



097856907

## MINISTERO DELL'INDUSTRIA, DEL COMMERCIO E DELL'ARTIGIANATO

DIREZIONE GENERALE DELLA PRODUZIONE INDUSTRIALE  
UFFICIO ITALIANO BREVETTI E MARCHI

REC'D 29 DEC 1999

WIPO

PCT



EP99/8856

INV. IND.

Autenticazione di copia di documenti relativi alla domanda di brevetto per

N. MI98A002509

*Si dichiara che l'unita copia è conforme ai documenti originali  
depositati con la domanda di brevetto sopraspecificata, i cui dati  
risultano dall'accluso processo verbale di deposito*

**PRIORITY DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

Roma, lì 2 NOV. 1999

IL REGGENTE  
IL DIRETTORE DELLA DIVISIONE

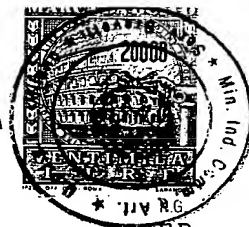
Elena Padellaro DI CINTIO

## AL MINISTERO DELL'INDUSTRIA DEL COMMERCIO E DELL'ARTIGIANATO

MODULO A

UFFICIO ITALIANO BREVETTI E MARCHI - ROMA

DOMANDA DI BREVETTO PER INVENZIONE INDUSTRIALE, DEPOSITO RISERVE, ANTICIPATA ACCESSIBILITÀ AL PUBBLICO



## A. RICHIEDENTE (I)

1) Denominazione S.I.S.S.A. SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI ED  
 Residenza TRIESTE codice 80035060328  
 2) Denominazione \_\_\_\_\_  
 Residenza \_\_\_\_\_ codice \_\_\_\_\_

## B. RAPPRESENTANTE DEL RICHIEDENTE PRESSO L'U.I.B.M.

cognome nome Dr. Diego Pallini ed altri cod. fiscale \_\_\_\_\_  
 denominazione studio di appartenenza Notarbartolo & Gervasi S.p.A.  
 via C.so di Porta Vittoria n. 9 città Milano cap 20122 (prov) MI

## C. DOMICILIO ELETTIVO destinatario

via \_\_\_\_\_ n. \_\_\_\_\_ città \_\_\_\_\_ cap \_\_\_\_\_ (prov) \_\_\_\_\_

## D. TITOLO

classe proposta (sez/cl/sci) C12N gruppo/sottogruppo 15

Processo per la preparazione di genoteche, di polipeptidi utilizzando dette  
genoteche e i polipeptidi ottenuti.

## ANTICIPATA ACCESSIBILITÀ AL PUBBLICO:

SI ☐ NO ☐

SE ISTANZA: DATA

N° PROTOCOLLO

## E. INVENTORI DESIGNATI

cognome nome

cognome nome

1) Andrew Raymon Morton Bradbury 3) \_\_\_\_\_  
 2) Daniele Sblattero 4) \_\_\_\_\_

## F. PRIORITÀ

nazione o organizzazione

tipo di priorità

numero di domanda

data di deposito

allegato  
S/R

SCIOGLIMENTO RISERVE

Data

N° Protocollo

1) Nessuna \_\_\_\_\_  
 2) \_\_\_\_\_

## G. CENTRO ABILITATO DI RACCOLTA CULTURE DI MICROORGANISMI, denominazione

## H. ANNOTAZIONI SPECIALI

Nessuna



## DOCUMENTAZIONE ALLEGATA

N. ps.

Doc. 1) ☐ PROV n. pag. 67  
 Doc. 2) ☐ PROV n. tav. 11  
 Doc. 3) ☐ RIS  
 Doc. 4) ☐ RIS  
 Doc. 5) ☐ RIS  
 Doc. 6) ☐ RIS  
 Doc. 7) ☐

riassunto con disegno principale, descrizione e rivendicazioni (obbligatorio 1 esemplare).....

disegno (obbligatorio se citato in descrizione, 1 esemplare).....

lettera d'incarico, procura o riferimento procura generale .....

designazione inventore .....

documenti di priorità con traduzione in italiano .....

autorizzazione o atto di cessione .....

nominativo completo del richiedente .....

SCIOGLIMENTO RISERVE

Data

N° Protocollo

confronta singole priorità

8) attestati di versamento, totale lire

Novecentoquindicimila.=

COMPILATO IL 19/11/1998

FIRMA DEL(I) RICHIEDENTE(I)

Diego Pallini

obbligatorio

CONTINUA SI/NO ☒ NODEL PRESENTE ATTO SI RICHIEDE COPIA AUTENTICA SI/NO ☒ SI

UFFICIO PROVINCIALE IND. COMM. ART. DI

MILANO

codice 15

VERBALE DI DEPOSITO

NUMERO DI DOMANDA

MI98A 002509

Reg. A.

L'anno millenovecento

NOVANTOTTO

il giorno

DICIANNOVE

del mese di

NOVEMBRE

il(i) richiedente(i) sopraindicato(i) ha(hanno) presentato a me sottoscritto la presente domanda, corredata di n.

00 fogli aggiuntivi per la concessione del brevetto sopraportato.

## I. ANNOTAZIONI VARIE DELL'UFFICIALE ROGANTE

IL DEPOSITANTE

timbro

L'UFFICIALE ROGANTE

RIASSUNTO INVENZIONE CON DISEGNO PRINCIPALE, DESCRIZIONE E RIVENDICAZIONE

NUMERO DOMANDA

M198A00239

REG. A

DATA DI DEPOSITO

16/11/1998

NUMERO BREVETTO

DATA DI RILASCIO

/ /

## D. TITOLO

Processo per la preparazione di genoteche, di polipeptidi utilizzando dette genoteche e i polipeptidi ottenuti.

## L. RIASSUNTO

E' descritto un processo per la generazione di diversità genica mediante la preparazione di genoteche e all'utilizzo di dette genoteche per la preparazione di nuovi polipeptidi e ai polipeptidi così ottenuti, per esempio sbp o scFv con migliorata efficienza.

Il processo per la preparazione di una genoteca comprende: a) l'introduzione di molecole di acido nucleico o di particelle infettanti comprendenti dette molecole di acido nucleico all'interno di una popolazione di cellule. Queste molecole di acido nucleico comprendono due parti: I) una sequenza di acido nucleico identica per ogni molecola; e II) una sequenza di acido nucleico variabile, non necessariamente continua, la quale è il substrato per la ricombinazione; e b) ricombinazione avviene tra le molecole di acido nucleico inserite intracellularmente durante la fase a).

## M. DISEGNO



Domanda di brevetto per Invenzione Industriale dal Titolo:

"Processo per la preparazione di genoteche, di polipeptidi utilizzando dette genoteche e i polipeptidi ottenuti."

Titolare: S.I.S.S.A. - SCUOLA INTERNAZIONALE SUPERIORE DI  
STUDI AVANZATI

con sede in: TRIESTE

MI 9 8 A 0 0 2 5 0 9

Inventori designati: BRADBURY Andrew Raymon Morton;

SBLATTERO Daniele

19 NOV. 1998

Depositata il

con il N.

\* \* \* \* \*

### **Campo dell'Invenzione**

La presente invenzione si riferisce ad un processo per la generazione di diversità genica mediante la preparazione di genoteche e all'utilizzo di dette genoteche per la preparazione di nuovi polipeptidi e ai polipeptidi così ottenuti.

### **Stato della tecnica**

La creazione della diversità in una popolazione di polipeptidi è diventata un importante strumento per l'ottenimento di polipeptidi derivati con caratteristiche utili. Questo richiede la disponibilità di metodi efficienti per la creazione di diversità congiuntamente a metodi per la selezione o screening di elementi aventi le caratteristiche desiderate.

La creazione della diversità può essere realizzata mediante metodi diversi, tra cui:

- i) mutagenesi casuale (random) - delle mutazioni vengono inserite casualmente in una sequenza di DNA; ciò può essere realizzato

mediante diversi metodi tra cui la mutagenesi chimica, l'utilizzo di ceppi batterici mutatori e PCR incline all'errore (error prone PCR);

ii) mutagenesi ricombinatoriale - si determina una ricombinazione tra catene di DNA differenti tale che le mutazioni, differenze o geni che sono originalmente presenti su catene diverse vengono a trovarsi insieme sulla stessa catena di DNA;

iii) mutagenesi diretta - le mutazioni vengono inserite in una sequenza di DNA codificante per una proteina in siti corrispondenti a quella porzione della proteina che è conosciuta o si pensa partecipi a particolari processi, come attività di legame o attività enzimatiche; ciò può essere realizzato mediante mutagenesi sito-specifica o per PCR.

Di queste tre forme di mutagenesi, la seconda (ricombinatoriale) e la terza (diretta) sembrano essere le più efficienti. Tuttavia, la mutagenesi diretta richiede una dettagliata conoscenza strutturale e non può essere applicata con modalità casuale (random). Gli altri due metodi (casuale e ricombinatoriale) possono invece essere applicati con modalità casuale e i risultati possono essere direttamente paragonati tra loro.

La mutagenesi ricombinatoriale per lo scopo di creare genoteche è stata eseguita solamente in vitro mediante la tecnica della ricombinazione generale (DNA shuffling) (Stemmer, 1994, *Proc Natl Acad Sci, USA* 91:10747-10751) ma mai intracellularmente. Mediante shuffling, clonaggio e trasfezione di DNA, sono state modificate le proprietà di un diverso numero di proteine, tra cui, la beta lattamasi, la galattosidasi ed anticorpi. Utilizzando la tecnica del DNA shuffling su geni strettamente correlati, piuttosto che mutazioni accumulate con la PCR, sono state

prodotte delle nuove proteine che presentano delle proprietà non presenti in nessuna delle proteine codificate dai geni utilizzati come substrati per la tecnica shuffling. Esempi comprendono gli interferoni e le cefalosporine (Cramer A. et al., 1998, DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 391:288-91).

Una particolare applicazione in cui la diversità viene creata dall'assortimento combinatoriale di due differenti geni è la creazione di una genoteca fagica combinatoriale di anticorpi i quali vengono espressi sulla superficie di fagi filamentosi o fagemidi (phage display system) (Marks J.D. et al., 1991. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.* 222:581-597).

Il dominio di legame degli anticorpi è codificato dalle regioni V, e in particolare da due distinti domini VH e VL. E' stato possibile separare i domini V dal resto della immunoglobulina e creare nuovi polipeptidi leganti funzionanti. I frammenti Fv sono costituiti da VH e VL tenuti assieme da legami non covalenti. Tuttavia, essi sono relativamente instabili e le due catene si separano facilmente. Legando le due catene con un polipeptide linker si ottiene un frammento a singola catena (scFv) che è molto più stabile. Il frammento Fab consiste di VH e CH1 legati covalentemente attraverso legami di solfuro tra VL e CL. Questo frammento risulta stabile ma presenta lo svantaggio di essere particolarmente difficile da preparare in batteri ed ha la tendenza ad aggregarsi. La maggior parte delle genoteche di anticorpi costruite su fagi filamentosi sono nella forma scFv ed alcune sono state preparate nella forma Fab.

Sebbene anticorpi contro una grande quantità di antigeni siano stati isolati usando una genoteca di anticorpi su fagi o fagemidi (phage antibody library) (per esempio, Griffiths A.D. et al., 1994. Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* 13:3245-3260 e Marks et al., 1991), la procedura non è ancora largamente utilizzata. Ciò è dovuto alla difficoltà di preparare delle ampie genoteche di anticorpi su fagi o fagemidi, che richiedono che la genoteca di almeno  $10^9$  cloni indipendenti sia creata da una buona fonte di geni VH e VL diversi. In generale, tali genoteche vengono preparate eseguendo un grande numero di ligazioni e trasfezioni, e una volta avvenute presentano il limite di non poter essere amplificate senza contemporaneamente portare ad una riduzione della diversità. In generale, l'affinità degli anticorpi isolati è proporzionale alla dimensione iniziale della genoteca usata per la selezione.

La ricombinazione a siti specifici è stata proposta come un metodo alternativo al clonaggio per la preparazione di vaste genoteche, e per la preparazione di anticorpi con nuove specificità. Per ricombinare le catene VH e VL nei Fab (Griffiths et al. 1994) o nei scFv (Tsurushita N. et al., 1996. Phage display vectors for in vivo recombination of immunoglobulin heavy and light chain genes to make large combinatorial libraries. *Gene* 172:59-63) è stata usata la cre ricombinasi mentre in un altro caso viene usata la lambda ricombinasi per ricombinare Fab (Geoffroy F. et al., 1994. A new phage display system to construct multicombinatorial libraries of very large antibody repertoires, *Gene* 151:109-113). Le due ricombinasi usate hanno differenti proprietà: la cre



ricombinasi catalizza la ricombinazione reversibile tra siti lox di 34 bp identici mentre la lambda ricombinasi catalizza la ricombinazione irreversibile nelle condizioni usate tra una sequenza lunga attP che si trova nel genoma di lambda ed una sequenza corta attB che si trova nel genoma del batterio.

Mentre la ricombinazione tra singoli anticorpi per ricostituire le regioni di legame funzionali è stata dimostrata per i casi sopra citati, una genoteca di anticorpi su fagi è stata preparato solamente utilizzando la cre ricombinasi per ricombinare i Fab (quindi ricombinazione per siti specifici) ed è stata descritta in WO 93/19172 e Griffiths et al, 1994.

In entrambe queste pubblicazioni viene preparata una genoteca determinando la diversità mediante ricombinazione per siti specifici di un primo vettore contenente il gene per la catena VH e un secondo vettore contenente il gene per la catena VL, con la produzione risultante di vettori di combinazione contenenti entrambi i geni codificanti per VH e VL. Tali primo e secondo vettore, diversi tra loro, possono essere un plasmide, un fago o un fagemide.

La permanenza in una cellula, da parte di vettori di espressione diversi e portanti i geni da ricombinare, si riferisce ad uno dei problemi dello stato dell'arte che necessitano di essere superati per poter creare diversità utilizzando il processo di ricombinazione indotta nelle cellule. Tale problema può essere individuato nel come i diversi DNA che sono substrato per la ricombinazione sono introdotti nella cellula in cui avviene tale ricombinazione, e inoltre in che modo i diversi substrati di DNA per la ricombinazione vengono mantenuti all'interno della cellula



per un tempo sufficientemente lungo per far avvenire la ricombinazione.

I documenti WO 93/19172 e Griffiths et al, 1994, sopra riportati, utilizzano vettori diversi tra loro per poter introdurre le due genoteche di VH e VL dentro la stessa cellula. Un plasmide che non contiene gene 3 è trasfettato mentre il secondo, un fago, infetta la cellula. Questo perché è stato riportato che batteri contenenti un fago filamentoso Ff sono resistenti a successive infezioni di altri fagi Ff (Boeke J.D. et al., 1982.

Effects of bacteriophage f1 gene III protein on the host cell membrane.

*Mol Gen Genet* 186:185-92) e l'infezione è il modo più efficiente di introdurre DNA in una cellula. Questa resistenza deriva dall'inibizione della formazione del pilus F' che è il recettore sulla superficie cellulare batterica utilizzato dal fago Ff per entrare nel batterio. Questa inibizione della sintesi del pilus F' è indotta dalla proteina del gene 3, che è presente sul fago ed è prodotta rapidamente in seguito all'infezione.

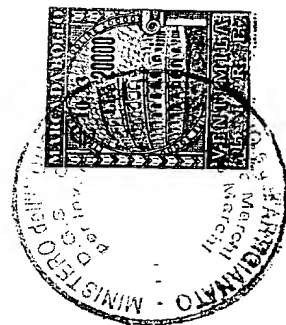
Per ovviare a questo problema e per poter mantenere il DNA substrato nella cellula per un tempo sufficientemente lungo per poter far avvenire la ricombinazione, si è cercato di mettere i substrati per la ricombinazione su differenti plasmidi contenenti le origini di replicazione appartenenti a gruppi con diversa incompatibilità e diverse resistenze per gli antibiotici. Questo poiché è noto agli esperti della tecnica che plasmidi che utilizzano lo stesso sistema di replicazione (origine di replicazione) non possono co-esistere stabilmente e sono detti essere incompatibili (pagine 1.3-4 del testo Sambrook J. Et al., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbour Laboratory Press). Per ogni plasmide e per ogni origine di

replicazione è anche richiesta una diversa resistenza agli antibiotici. Ciò limita il numero totale di plasmidi diversi che possono essere mantenuti all'interno di ogni singola cellula e conseguentemente il numero di substrati di ricombinazione, che possono essere presenti all'interno di ogni singola cellula, al numero di origini di replicazione diversi e alle resistenze per gli antibiotici che possono essere usati. Questa ristretta limitazione del numero dei potenziali substrati di ricombinazione richiede che i substrati di ricombinazione siano specificatamente designati. Ciò è appropriato quando si studiano i processi di ricombinazione, ma è inverosimile che si possa realizzare la mutagenesi ricombinatoriale casuale (DNA shuffling) come viene usata in vitro, per creare diversità a livello intracellulare quando il numero di substrati possibili per la ricombinazione è così limitato. Inoltre, anche se il numero di plasmidi diversi con origini di replicazione e resistenze agli antibiotici diverse possano essere introdotti in una singola cellula, la ricombinazione tra detti plasmidi creerebbe dei nuovi plasmidi con miscele sia per l'origine di replicazione che per la resistenza agli antibiotici che li renderebbe estremamente difficili da analizzare ed utilizzare.

Un caso in cui la stessa origine di replicazione è presente su due vettori diversi presenti simultaneamente all'interno di una cellula è quando si producono i fagemidi. Durante questa procedura i batteri vengono infettati o trasfettati con DNA di fagemide e le particelle (fagemidiche) infettanti vengono prodotte con la successiva infezione di un fago *helper*. Sia il fagemide che il fago *helper* contengono l'origine di

replicazione Ff. Tuttavia, il fagemide contiene sia l'origine di replicazione plasmidica (per esempio pUC) che l'origine di replicazione Ff (per esempio M13), mentre il fago helper può anche contenere un'origine di replicazione plasmidica di un diverso gruppo di incompatibilità rispetto al fagemide (per esempio p15A in caso di M13K07, uno dei fagi *helper* più comunemente utilizzati) in aggiunta all'origine Ff. La replicazione del DNA fagemidico avviene grazie all'origine di replicazione plasmidica (pUC) e non all'origine Ff che è invece usata per impacchettare il fagemide o il fago *helper* in particelle infettanti (questa origine ha la duplice funzione di impacchettare le particelle e di replicare il DNA), mentre la replicazione del DNA del fago *helper* avviene grazie sia all'origine di replicazione del plasmide (p15A) in M13K07 che all'origine di replicazione Ff (vedi pagine 4.17-20 di Sambrook et al., 1989). In entrambi i casi, vengono usate diverse origini di replicazione per il DNA del fagemide e quello del fago *helper*.

Come esemplificativo dello stato dell'arte, per maggiore chiarezza, viene qui illustrato in Figura 1, in forma schematica semplificata, il metodo descritto in Griffiths et al, 1994. In Fig.1 è mostrata la ricombinazione per siti specifici che avviene all'interno di una singola cellula batterica la quale viene prima trafettata da un plasmide contenente il gene per la catena pesante, VHA, e quindi infettata da un fago portante il gene per la catena leggera, VLA. Dopo la ricombinazione si ottiene la combinazione funzionale VHA/VLA, così come le tre forme non funzionali, come illustrato nella Figura 1. Per ottenere questa genoteca ricombinata, vengono preparate due genoteche sintetiche primarie: 1)

*DP*

una genoteca VH di  $10^8$  geni VH+CH1 clonati in un vettore plasmidico con il dominio VH+CH1 fiancheggiato da due segnali di ricombinazione (loxP wild type e loxP 511); 2) una genoteca VL di  $8 \times 10^5$  di geni della catena leggera VL+CL preparati in un fago fd contenente un gene VH+CH1 unico quindi con una unica specificità (*dummy*) fiancheggiato dagli stessi segnali di ricombinazione.

La genoteca viene quindi preparata infettando batteri contenenti il plasmide avente la genoteca VH ( introdotta mediante trasfezione) con fagi aventi la genoteca VL. I segnali di ricombinazione sono capaci di ricombinare solamente in maniera omologa, cioè loxP wt con loxP wt e loxP 511 con loxP511, e come risultato la catena VH dummy può essere scambiato con la genoteca VH, come illustrato in Figura 1, in modo da creare la genoteca. La diversità in tale genoteca è dipendente dal numero di cellule infettate, e non può superare questa cifra. Anticorpi con affinità nanomolari sono stati isolati da questa genoteca. Come ricordato sopra, è possibile la ricombinazione tra il gene della catena pesante singola e il gene della catena leggera singola, ma non è possibile la ricombinazione tra geni V derivati da più di un plasmide ed un fago, cioè tra più di una coppia di geni V.

Questo metodo presenta diversi problemi:

- a) il sistema di ricombinazione a siti specifici è reversibile e la genoteca viene quindi contaminata con il fago di partenza (contenente le catene VH dummy) e con i plasmidi (contenenti la genoteca VH+CH1 o il dummy VH+CH1) a livelli relativamente alti, come mostrato in Figura 1;

- b) si creano dei prodotti secondari del processo di ricombinazione che consiste di plasmidi doppi (non mostrato in Figura 1);
- c) è stato trovato che molti plasmidi, come quelli utilizzati in tale metodo basati sul plasmide pUC19, sebbene non contengano l'origine di replicazione Ff possono comunque essere incorporati nelle particelle fagiche con efficienze diverse; ciò fa sì che la genoteca finale contenga una miscela di molti elementi genetici differenti riducendo così la diversità funzionale effettiva;
- d) la ricombinazione avviene tra un singolo fago a DNA e una singola molecola plasmidica di DNA, mentre la ricombinazione tra più plasmidi diversi non è possibile; la conclusione è che i batteri producono un singolo nuovo anticorpo ricombinato e questo limita la dimensione della genoteca ottenibile al numero dei batteri usati nella fase di ricombinazione, il quale in pratica non va oltre il limite di  $5 \times 10^{12}$  per 10 litri; per aumentare le dimensioni della genoteca deve essere usato un maggior numero di batteri;
- e) i prodotti di ricombinazione costituiscono un prodotto finale e non è possibile effettuare una ulteriore ricombinazione tra: a) le catene VH e le catene VL ricombinate; b) le VH ricombinate e la genoteca di catene VL. L'unica possibilità è quella di ricombinare le catene VL ricombinate con la genoteca di catene VH originarie.

In conclusione, esiste l'esigenza in questo campo di applicazione di un metodo efficiente per la generazione di diversità di acidi nucleici e la creazione di genoteche da utilizzare per la preparazione di nuovi polipeptidi derivati che presentino delle caratteristiche utili.

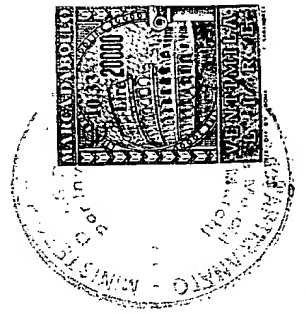
I metodi sino ad ora conosciuti, come quello descritto in WO 93/19172 o in Griffiths et al, 1994, presentano dei seri inconvenienti di limitazione delle dimensioni della genoteca che è proporzionale al numero di cellule batteriche utilizzate per la fase di ricombinazione.

Il problema tecnico dei metodi conosciuti nello stato dell'arte è legato alla difficoltà, del metodo sino ad ora adottato, di introdurre nella cellula l'acido nucleico substrato della ricombinazione e alla difficoltà di mantenere all'interno della cellula l'acido nucleico substrato per un tempo sufficientemente lungo per far avvenire la ricombinazione.

### **Sommario dell'Invenzione**

Gli autori della presente invenzione hanno risolto questi problemi tecnici e hanno realizzato un metodo per la creazione di vaste genoteche ad alta diversità e in cui la diversità creata in una singola cellula batterica è notevolmente superiore a quella che si ottiene con i processi dello stato dell'arte ed hanno trovato che, sorprendentemente, più di una molecola di acido nucleico costituita da due parti: I) una sequenza di acido nucleico identica per ogni molecola; e II) una sequenza di acido nucleico variabile, non necessariamente contigua, la quale è il substrato per la ricombinazione, possono essere introdotte all'interno di almeno una cellula di una popolazione di cellule.

Una caratteristica della presente invenzione consiste pertanto in un metodo per la creazione di genoteche ad elevata diversità in cui: a) almeno due membri di una popolazione di molecole di acido nucleico costituite da due parti: I) una sequenza di acido nucleico identica per ogni molecola, e II) una sequenza di acido nucleico variabile, non



necessariamente contigua, la quale è il substrato per la ricombinazione, vengono inserite all'interno di almeno una cellula di una popolazione di cellule; e b) avviene una ricombinazione tra le molecole di acido nucleico inserite.

Secondo una ulteriore caratteristica, la presente invenzione consiste in un metodo per la creazione di genoteche ad elevata diversità mediante l'infezione di almeno una cellula di una popolazione di cellule con almeno due particelle infettive comprendenti le molecole di acido nucleico secondo l'invenzione.

Un'altra caratteristica riguarda un processo di introduzione di molecole di acido nucleico secondo l'invenzione in almeno una cellula di una popolazione di cellule o l'infezione di detta cellula con più particelle infettive comprendenti dette molecole di acido nucleico ed in cui la ricombinazione avviene per siti specifici o per ricombinazione generale.

L'invenzione riguarda inoltre dette genoteche di una molteplicità di particelle infettive o rgdp che possono esprimere sulla loro superficie i polipeptidi ed inoltre un processo utilizzante dette genoteche per la preparazione di polipeptidi.

Una ulteriore caratteristica riguarda la preparazione di uno dei due membri di una coppia di elementi aventi tra loro legami specifici (sbp) ed ancora un'altra caratteristica riguarda la preparazione di anticorpi o di suoi frammenti.

L'invenzione si riferisce inoltre ai prodotti, tra cui, genoteche secondarie, particelle infettanti, rgdp, polipeptidi o loro frammenti, rbp, anticorpi e cellula infettate e polylinker come rivendicato.

### Descrizione delle Figure

**Figura 1:** Ricombinazione con sistema a due vettori (secondo Griffiths et al, 1994)

Sono descritti il plasmide iniziale (contenente la genoteca VH+CH1) ed il fago iniziale (contenente la genoteca VL+CL). E' inoltre illustrata la situazione all'interno di un singolo batterio, contenente un plasmide avente il gene della catena VHA, che viene infettato con un fago portante il gene per la catena VLA. Soltanto uno dei prodotti che contiene VHA e VLA è funzionale. La genoteca finale è prodotta quando la ricombinazione descritta avviene in molte cellule e la sua dimensione è proporzionale al numero di cellule usate.

**Figura 2:** Ricombinazione tra siti specifici in un sistema ad un solo (tipo di) vettore (schema generico secondo la presente invenzione)

Viene mostrata l'infezione di un batterio da parte di due fagemidi contenente le regioni VHA/VLA e VHB/VLB. Dopo la ricombinazione, quattro prodotti risultano presenti: gli originali VHA/VLA e VHB/VLB e i due nuovi prodotti VHA/VLB e VHB/VLA. Tutti i fagemidi, sia prima che dopo la ricombinazione, contengono anticorpi funzionali (scFv). Oltre al numero di cellule usato, la dimensione della genoteca finale è anche determinata dal numero di fagemidi che entrano in una cellula e quanto è estesa la ricombinazione che vi avviene.

**Figura 3:** Sequenze del linker loxP 511 del scFv

Sono indicate le sequenze di DNA e le sequenze aminoacidiche dedotte del linker loxP 511 usato per la creazione di genoteche primarie e in generale per il clonaggio di scFv. I siti di restrizione usati per il clonaggio



delle regioni V sono indicate sotto le sequenze di DNA. La sequenza del sito loxP 511 coinvolta nella ricombinazione è riportata in *corsivo*, i siti di restrizione usati per il clonaggio sono sottolineati, e le ripetizioni invertite nel sito di ricombinazione loxP 511 sono doppiamente sottolineate. Le regioni V degli anticorpi sono clonati nell'ordine: VL- linker-VH.

#### **Figura 4: pDAN5 polylinker**

E' riportata la sequenza del polylinker pDAN5 per il clonaggio di scFv. Sono indicate le posizioni delle catene VL e VH (le sequenze sono omesse), e in neretto sono evidenziati i siti di restrizione usati per clonarle. Sono indicati, con doppia sottolineatura, i siti loxP 511 e loxP wild type (w.t.) usati per indurre la ricombinazione. Sono inoltre indicate altre caratteristiche importanti quali la sequenza leader, il marcatore SV5 usato per riconoscere il scFv, il marcatore His per la purificazione, il codone stop amber e il gene 3.

#### **Figura 5: Mappa di pDAN5**

E' mostrata la mappa del vettore display pDAN5 con il scFv D1.3 clonato. I siti usati per il clonaggio del gene V sono in neretto.

#### **Figura 6 (A e B): Sequenza di pDAN5**

E' riportata la sequenza del vettore display pDAN5 (e la sequenza nucleotidica di scFv D1.3 in esso clonato).

#### **Figura 7: Schema di esperimento di D1.3 ricombinante**

Due fagemidi contenenti VL/X-VH/D1.3 e VL/D1.3-VH/Y (dove X e Y rappresentano due anticorpi non conosciuti) sono utilizzati per infettare batteri esprimenti cre ricombinasi o batteri wild type DH5 $\alpha$ F (che non esprimono cre ricombinasi) ad alta molteplicità di infezione (MOI). Dopo

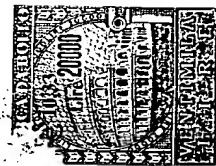
crescita per circa 12 ore, i fagemidi vengono preparati, re-infettati su cellule DH5 $\alpha$ F (quindi non più su cellule esprimenti cre) in modo da rendere genotipo e fenotipo accoppiati (coerenti tra loro), e testati con PCR ed ELISA per verificare la presenza del scFv VL/D1.3-VH/D1.3 funzionale.

**Figura 8:** Schema per la preparazione di una genoteca secondaria mediante ricombinazione intracellulare.

Batteri esprimenti cre ricombinasi sono infettati con una genoteca primaria di fagemidi ad un livello di alta molteplicità di infezione (MOI) pari a 200:1. Dopo la ricombinazione, i fagemidi sono preparati e re-infettati in cellule DH5 $\alpha$ F ad un livello di MOI minore di 1 in modo da accoppiare il genotipo con il fenotipo. I fagemidi prodotti da queste cellule con fenotipo e genotipo accoppiato vengono usati per la selezione della genoteca. In figura, per questione di semplicità schematica viene indicata una diversità creata dall'infezione di un singolo batterio con solo quattro fagemidi (con diversi geni VH e VL, in figura differenziati da una diversa retinatura). Comunque, secondo la presente invenzione, si è ottenuta una infezione di una singola cellula batterica con un numero molto più elevato di fagemidi. Test di PCR fingerprinting mostrati in figura 9 confermano questi dati.

**Figura 9:** Valutazione della diversità di batteri esprimenti cre ricombinasi infettati da una molteplicità di fagemidi

Colonie batteriche individuali esprimenti singoli fagemidi (dopo re-infezione ad un valore di MOI minore di 1) sono state isolate da due colonie dopo ricombinazione come descritto in Figura 8. I geni di VH e



VL ottenuti da tali colonie sono stati amplificati usando degli oligonucleotidi specifici per VH e VL (Sblattero, D. and Bradbury, A. 1998. A definitive set of oligonucleotide primers for amplifying human V regions. *Immunotechnology* 3:271-8) e identificati mediante mappatura usando BstNI. Le diverse bande di mappatura che corrispondono a diversi geni VH sono state numerate orizzontalmente da 1 a 17 per cellula 1, e da 1 a 18 per cellula 2, mentre le bande di mappatura del gene VL sono state numerate verticalmente da 1 a 12 per cellula 1 e da 1 a 14 per cellula 2. I valori dei mappatura di VH e VL trovati in 37 batteri individuali per cellula 1 e 41 batteri individuali per cellula 2 sono stati analizzati e riassunti nella tabella, con un numero che indica il numero di volte che ogni combinazione VH/VL è stata trovata. Questi dati danno una stima del numero e delle diverse combinazioni di VH/VL che sono state presenti nelle cellule 1 e 2 originali.

#### **Figura 10**

**Schema della ricombinazione generale utilizzato per creare un gene resistente alla cefotaxime.**

Una popolazione batterica viene infettata da una genoteca primaria di fagemidi di beta lattamasi con un rapporto fagi:batteri maggiore di 10:1. Avviene una ricombinazione che porta le mutazioni che originariamente erano in due geni diversi nello stesso gene. Solamente i fagemidi contenenti il gene della beta lattamasi ricombinato sono in grado di conferire resistenza alla cefotaxime. La genoteca primaria utilizzata ha una diversità di  $3 \times 10^5$ . Tale diversità e' rappresentata in Figura da due fagemidi ognuno dei quali contiene una singola mutazione. Nessuna di

queste mutazioni da sole è in grado di conferire resistenza alla cefotaxime, la resistenza si ha solamente quando le due mutazioni sono presenti nello stesso gene (grazie alla ricombinazione).

### **Definizioni e Abbreviazioni**

Per una chiara comprensione vengono qui di seguito riportate le definizioni che verranno utilizzate della presente domanda.

- *Proteina di rivestimento (coat protein)* - una proteina che si trova sulla superficie di una particella infettiva che riveste l'acido nucleico in essa compreso; le proteine di rivestimento possono essere definite minori o maggiori a seconda del numero di copie presenti sulla particella infettiva; nel fago Ff, p8 è la maggiore proteina di rivestimento e presenta approssimativamente 2800 copie, mentre p3 è la minore proteina di rivestimento ed è presente in 3-5 copie;
- *infezione* - processo tramite il quale una particella infettiva entra in una cellula in virtù di una interazione specifica tra una proteina di rivestimento o proteine che si trovano sulla superficie della particella infettiva ed un recettore sulla superficie cellulare; in caso del fago Ff o di un fagemide, l'interazione avviene tra p3 sulla superficie del fago e il pilus F del batterio;
- *particella infettiva* - una molecola di acido nucleico, rivestita con "proteine di rivestimento", capaci di entrare in una cellula con alta efficienza in virtù di interazioni specifiche tra la proteina o le proteine di rivestimento e un recettore sulla superficie cellulare; un esempio di particella infettiva è il fago o fagemide Ff;
- *genoteca primaria* - una collezione di acidi nucleici composta da

molecole individuali di acido nucleico ognuna delle quali è formata da due parti: una sequenza di acido nucleico identica per ogni molecola ed una parte con sequenza di acido nucleico variabile, non necessariamente contigua, la quale è il substrato per la ricombinazione. Ogni molecola di acido nucleico differisce dalle altre molecole di acido nucleico trovati nella genoteca;

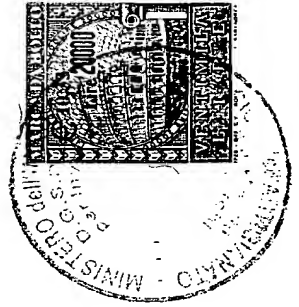
- *genoteca secondaria*: una collezione di acidi nucleici composta da molecole individuali di acido nucleico derivati da una genoteca primaria nella quale la diversità è stata aumentata tramite il processo di ricombinazione, sia ricombinazione generale, sia ricombinazione a siti specifici;
- *genoteca di particelle infettive (library of infectious particles)* - una raccolta di particelle infettive ognuna delle quali contiene una molecola di acido nucleico che differisce dalle molecole di acido nucleico contenute dalle altre particelle infettive presenti nella genoteca; queste differenze sono in genere ristrette ad un gene specifico codificante una specifica funzione le cui proprietà si intende modificare, mentre le altre parti della particella infettiva, in genere, rimangono inalterate; quando il numero delle particelle infettive usato è superiore alla diversità della genoteca ciascuna di queste particelle potrà essere presente in più di una copia;
- *substrato di ricombinazione*: una molecola di acido nucleico la quale è in grado di ricombinare con un altro substrato di ricombinazione, sia per la presenza di sequenze specifiche di ricombinazione (per esempio loxP, dove la ricombinazione è catalizzata da la ricombinasi

specifica, cre ricombinasi), che per la presenza di similarità nelle sequenze nucleotidiche che permettono l'azione del sistema cellulare di ricombinazione (ricombinazione generale);

- *ricombinazione specifica* ricombinazione che avviene a siti specifici (per esempio loxP) riconosciuti dalla ricombinasi (per esempio la cre ricombinasi) che agisce in queste sequenze;
- *molteplicità di infezione (multiplicity of infection) (MOI)* - il rapporto tra il numero di fagi o fagemidi (misurato come delle unità formanti delle colonie o unità formanti delle placche) aggiunti ad una coltura di cellule ed il numero delle cellule; un MOI di 200 indica che i fagi o fagemidi sono 200 volte in eccesso rispetto al numero delle cellule;
- *replicable genetic display package (rgdp)* - una particella biologica avente informazione biologica che fornisce alla particella la capacità di replicare in appropriate condizioni; la particella può esprimere alla sua superficie almeno una parte di un polipeptide; il polipeptide può essere codificato dall'informazione genetica nativa della particella o da informazioni artificiali posizionate nella particella; il polipeptide o una sua parte o frammento espressa sulla particella può essere qualsiasi elemento di sbp, per esempio un scFv, domini di catene pesanti (H) o leggere (L) di una molecola immunoglobulinica, un enzima, un recettore, ecc.;
- *vettore*: una molecola di DNA che è in grado di replicarsi, e qualsiasi altra sequenza in essa contenuta, in una cellula batterica. Normalmente, i vettori contengono un origine di replicazione, una resistenza contro un antibiotico ed alcuni siti di restrizione che

permettono il clonaggio di altri pezzi di DNA;

- *vettore d'esposizione*: un vettore usato nella creazione dei rgdp. Normalmente i vettori d'esposizione, oltre all'origine di replicazione e resistenza antibiotica, contengono un gene che codifica per una proteina, nella quale può essere espresso un polipeptide normalmente estraneo alla particella biologica, che viene esposto all'esterno di una particella biologica. Contengono, inoltre, un'altra origine di replicazione che permette l'incorporazione del DNA del vettore nella particella biologica che espone il polipeptide estraneo. pDAN5, un vettore fagemidico, è un esempio di un vettore d'esposizione;
- *fagemide (phagemid)* - una particella infettiva che consiste di un plasmide di DNA contenente l'origine di replicazione di fago filamentoso che è rivestito da proteine di rivestimento del fago filamentoso fornite dal fago *helper*, ugualmente il termine fagemide può anche riferirsi al vettore di DNA contenuto nella particella infettiva; un fagemide Ff si riferisce ad un fagemide che usa l'origine di replicazione della famiglia dei fagi filamentosi Ff ed ha bisogno di un fago *helper* derivato dalla stessa famiglia Ff, pDAN5 è un esempio di un fagemide di tipo Ff;
- *frammenti di anticorpi specifici* - frammenti di anticorpi che sono capaci di legare antigeni specifici; esempi di frammenti di anticorpi comprendono scFv, Fv, Fab;
- *linker* - una sequenza di amino acidi che si trova tra la fine di VL e l'inizio di VH legando covalentemente le due catene, permettendo di



formare una corretta struttura scFv. In generale un linker è una sequenza amino acidica che lega due polipeptidi covalentemente in modo formare una unica catena polipeptidica.

- *coppi di elementi aventi tra di loro un legame specifico (specific binding pair o sbp)* - una coppia di molecole (ciascuna essendo un elemento o membro di una coppia di elementi aventi tra di loro un legame specifico) che sono derivate naturalmente o prodotto per via sintetica, e sono in grado di legarsi l'una all'altra specificamente; esempi di tipi di sbp comprendono antigene-anticorpo, biotina-avidina, ormone-recettore ormonale, recettore-ligando, enzima-substrato, igG-proteina A, ecc,;
- *regioni V* - frammenti che comprendono la regione V (variabile) delle immunoglobuline.

#### **Descrizione dettagliata dell'Invenzione**

Con il metodo secondo la presente invenzione, almeno una cellula di una popolazione di cellule viene trasfettata con più di una molecola di acido nucleico, o infettata da più copie di una particella infettante comprendenti dette molecole di acido nucleico e aventi stessa origine di replicazione e resistenza per gli antibiotici, ciascuna di detta molecola di acido nucleico avente un diverso contenuto di substrato di acido nucleico per la ricombinazione. Con questo metodo la ricombinazione avviene anche se le molecole di acido nucleico o le particelle portanti l'acido nucleico da ricombinare hanno identica origine di replicazione e resistenza per gli antibiotici. Si è trovato che un numero molto alto di molecole di acido nucleico o particelle infettanti possono essere



introdotte in una singola cellula aumentando così enormemente la quantità di substrato per la ricombinazione. In conclusione, la diversità creata da una singola cellula è di molte volte superiore di quella riportata in WO 93/19172. Si è infatti calcolato che si può avere una diversità, in una singola cellula, fino 324 volte. Sulla base delle cifre riportate nella Figura 9, si può stimare che la diversità creata dentro una cellula ha un valore minimo di 29 (il numero delle diverse catene identificate) ed un valore massimo di 324 (il numero delle catene prodotto se la ricombinazione tra 18 diversi scFv è completa).

La presente invenzione si riferisce quindi ad un processo per preparare una genoteca e a detta genoteca, all'uso di detta genoteca per la preparazione di polipeptidi e a detti polipeptidi.

Per la preparazione di una genoteca secondo l'invenzione, si procede preparando una genoteca primaria, cioè una molteplicità di molecole di acido nucleico comprendenti due parti: I) una sequenza di acido nucleico identica per ogni molecola, e II) una sequenza di acido nucleico variabile, non necessariamente contigua, la quale è il substrato per la ricombinazione, oppure preparando una molteplicità particelle infettanti, ciascuna particella comprendente una molecola di acido nucleico secondo l'invenzione (ciascuna avente identica origine di replicazione e resistenza per gli antibiotici). Quindi si trasfetta o infetta una cultura cellulare con la genoteca primaria (cioè la molteplicità di molecole di acido nucleico o di particelle infettanti) con un rapporto molecole di acido nucleico o di particelle infettanti:cellule maggiore di 2, in modo che più di una molecola di acido nucleico entri in una cellula; in questa fase

avviene la ricombinazione intracellulare dell'acido nucleico substrato;

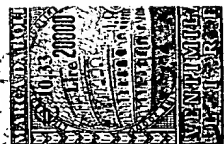
Secondo una caratteristica, la presente invenzione si riferisce dunque ad un processo per la preparazione di una genoteca (secondaria), comprendente le seguenti fasi:

- a) introduzione di almeno due membri di una popolazione di molecole di acido nucleico all'interno di almeno una cellula di una popolazione di cellule, detta popolazione di molecole di acido nucleico comprendente molecole individuali di acido nucleico ognuna delle quali essendo costituita da due parti: I) una sequenza di acido nucleico identica per ogni molecola; e II) una sequenza di acido nucleico variabile, non necessariamente contiguo, la quale è il substrato per la ricombinazione;
- b) la ricombinazione avviene tra le molecole di acido nucleico substrato inserite intracellularmente durante la fase a).

Secondo un'altra realizzazione, l'invenzione si riferisce ad un processo per la preparazione di una genoteca (secondaria) comprendente le seguenti fasi:

- a) almeno una cellula di una popolazione cellulare viene infettata con almeno due elementi di una popolazione di particelle infettanti, dove ciascuna di dette particelle comprende una molecola di acido nucleico secondo l'invenzione e come sopra definito;
- b) ricombinazione avviene tra l'acido nucleico substrato inserito nelle cellule durante la fase di infezione a).

Ovviamente con acido nucleico si intende DNA o RNA. Le cellule infettate possono essere cellule batteriche, preferibilmente *Escherichia coli*, o anche cellule eucariotiche. Le particelle infettive o le "replicable



genetic display package" (rgdp) possono essere qualsiasi particella infettiva che è in grado di infettare una cellula senza indurre la lisi, per esempio, fagi che non inducono la lisi, ad esempio fagi filamentosi (ad esempio fagi filamentosi della famiglia Ff), o fagemidi.

La ricombinazione della fase b) può avvenire sia tramite i sistemi ricombinatoriali presenti nella cellula ed in questo caso si parlerà di ricombinazione generale oppure può avvenire tra siti specifici.

L'efficienza della ricombinazione generale può essere aumentata usando ceppi batterici contenenti certe mutazioni, per esempio, mutS o recO98 oppure con trattamenti fisici, per esempio trattamenti UV.

Le molecole di acido nucleico per la trasfezione o le particelle infettanti dello stesso tipo, fagi o fagemidi, con i quali viene infettata una cellula, hanno in comune una sequenza identica (vettore *backbone*), che nel caso del vettore pDAN5 è costituita dall'origine di replicazione, la resistenza per ampicillina, un leader per la secrezione, i marcatori (SV5 e His), gene3 e l'origine di replicazione per Ff e allo stesso tempo differiscono l'una dall'altra per il diverso contenuto di molecole di acido nucleico substrato per la ricombinazione. Detta parte di acido nucleico, substrato per la ricombinazione, contiene le sequenze di acido nucleico che si intende ricombinare. Questi possono essere uno o più geni codificanti per uno o più polipeptidi, oppure sequenze di acido nucleico coinvolte in un attività biologica, per esempio sequenze per la regolazione genica. Un esempio di particella infettante o vettore di infezione secondo la presente invenzione è riportato in Figura 5 (la sua sequenza è riportata in Figura 6) e si riferisce a pDAN5 D1.3. D 1.3 è un

esempio di un scFv ed è sostituibile.

Per quanto riguarda la ricombinazione a siti specifici, questa avviene in uno o più siti specifici che si trovano all'interno di dette molecole di acido nucleico ed è catalizzata da ricombinasi, specifiche per detti siti, espresse dalla cellula, mentre la ricombinazione generale avviene in un modo casuale lungo l'acido nucleico ed è mediata dai sistemi cellulari di ricombinazione.

Preferibilmente, la ricombinazione a siti specifici avviene tra molecole diverse di acido nucleico su siti specifici loxP, ottenibili da colifago P1, o da sequenze mutate o sequenze derivate da siti loxP, preferibilmente loxP 511. Questi siti loxP sono presenti in dette molecole di acido nucleico, e la ricombinazione è catalizzata dalla cre ricombinasi espressa dalla cellula.

Secondo una particolare realizzazione, nel processo secondo l'invenzione, l'acido nucleico introdotto nelle cellule e substrato per la ricombinazione comprende:


- i) una prima sequenza di ricombinazione sito specifica; e
- j) una seconda sequenza di ricombinazione sito specifica diversa dalla prima;

ed in cui la ricombinazione sito specifica avviene tra le prime sequenze di ricombinazione sito specifiche su diverse molecole di acido nucleico e tra le seconde sequenze di ricombinazione sito specifiche sulle stesse molecole di acido nucleico in cui avviene la ricombinazione al primo sito, ma non tra una prima sequenza di ricombinazione sito specifica ed una seconda sequenza di

ricombinazione sito specifica.

Preferibilmente, la prima sequenza di ricombinazione sito specifica è loxP (ottenibile dal colifago P1) e la seconda sequenza di ricombinazione sito specifica è un mutante della sequenza loxP, preferibilmente loxP 511, la cellula trasfettata o infettata è preferibilmente una cellula batterica e ancora più preferibilmente E.coli e le particelle infettanti sono fagi o fagemidi, preferibilmente fagemidi derivati da Ff e contengono DNA fagemidico.

In un'altra realizzazione dell'invenzione, la molecola di acido nucleico o la particella infettante comprendente detto acido nucleico secondo l'invenzione comprende: I) una sequenza di acido nucleico identica per ogni molecola che rappresenta la porzione costitutiva (identico in ogni molecola o è essenziale per la replicazione e sopravvivenza all'interno della cellula), detta sequenza viene anche indicata con il termine vettore "backbone"; e II) una sequenza di acido nucleico variabile, non necessariamente contigua, la quale rappresenta il substrato della ricombinazione comprendente geni codificanti una prima catena polipeptidica ed una seconda catena polipeptidica di sbp tale che con la ricombinazione si ottengono molecole di acidi nucleici diverse codificanti per nuove catene o nuove combinazioni di catene di primo e secondo polipeptide. Preferibilmente, la ricombinazione tra geni di detta prima e seconda catena di polipeptidi di sbp indotta a livello di una prima sequenza di ricombinazione sito specifica che si trova tra i geni di dette prima e seconda catene polipeptidiche e a livello di una seconda sequenza di ricombinazione sito specifica che si trova tra i geni di dette



seconda e prima catene polipeptidiche, o viceversa.

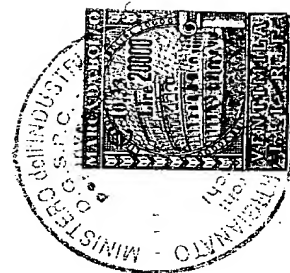
La genoteca secondaria secondo l'invenzione potrebbe essere quindi costituita da una cellula comprendente al suo interno acido nucleico da particelle infettanti o rdpg, e cioè fagi o fagemidi, e esprime sulla sua superficie (della cellula) polipeptidi o parti di essi.

La genoteca secondo l'invenzione può anche essere costituita da una molteplicità di particelle infettanti o rdgp esprimenti sulla loro superficie (legati a proteine di superficie delle particelle infettanti o rdgp) polipeptidi o frammenti di polipeptidi.

Questo tipo di genoteca o genoteca secondaria, quando esprime polipeptidi, avrà un contenuto genotipico dell'acido nucleico non accoppiato con il fenotipo dei polipeptidi o frammenti di polipeptidi espressi come è evidente dalla Figura 8. Il rapporto di infezione vettori fagici:cellule sarà preferibilmente di 200:1.

Tale genoteca secondaria può quindi essere utilizzata per una nuova infezione, ad un rapporto particella infettante o rgdp:cellula minore di 1, in modo che statisticamente ogni cellula sia generalmente infettata da una particella infettante o da un rgdp di un solo tipo per la preparazione di una seconda genoteca secondaria in cui il genotipo ed il fenotipo sono accoppiati (come risulta dalla Figura 8). Questa seconda genoteca può essere selezionata o sottoposta a screening a livello della cellula o rgdp derivati dalle cellule secondo il substrato di ricombinazione.

Secondo una ulteriore realizzazione, quindi, la presente invenzione si riferisce ad un processo per la preparazione di polipeptidi utilizzando la genoteca secondo l'invenzione.



La preparazione di polipeptidi comprende le seguenti fasi:

- a) viene preparata una genoteca secondaria secondo quanto descritto precedentemente, dove le cellule ricombinate esprimono uno o più polipeptidi;
- b) selezione o screening di uno o più polipeptidi.

Le fasi a) e b) possono essere ripetute ciclicamente secondo convenienza.

L'invenzione si riferisce quindi ad un processo per la preparazione di polipeptidi secondo l'invenzione, in cui detta genoteca secondaria, comprendente una molteplicità di particelle infettanti o rgdp in cui il fenotipo e il genotipo non sono accoppiati, è utilizzata per infettare cellule con un rapporto particelle infettanti o rgdp:cellule minore di 1 e viene prodotta una seconda genoteca secondaria comprendente una seconda molteplicità di particelle infettanti o rgdp aventi il fenotipo accoppiato al genotipo.

Nella fase a) di cui sopra, i polipeptidi, codificati da acidi nucleici, sono espressi sulla superficie di particelle infettanti. Secondo una realizzazione, la selezione o screening avviene tramite l'espressione di una attività codificata da una o più proteine. Secondo un'altra realizzazione, la selezione o screening avviene grazie ad un'attività biologica espressa all'interno di una cellula. Per esempio, una attività enzimatica, come quella selezionata per la degradazione di cefotaxime. Secondo un'ulteriore realizzazione, il polipeptide o i polipeptidi sono espressi in modo da fornire una popolazione di acidi nucleici codificanti uno o più polipeptidi selezionati o sottoposti a screening.

Detti polipeptidi, codificati da acidi nucleici presenti nelle particelle infettanti, possono essere espressi sulla superficie di "replicable genetic display package" (rgdp); e nella fase b) la selezione o screening del polipeptide o dei polipeptidi espressi avviene in modo da fornire una popolazione di acidi nucleici codificanti uno o più polipeptidi selezionati o sottoposti a screening.

Secondo una ulteriore realizzazione la fase a) del processo è ripetuta con una popolazione di acidi nucleici o con una popolazione di "replicable genetic display package" (rgdp) ottenuti dalla fase b).

La selezione o screening secondo l'invenzione può comprendere l'accoppiamento di polipeptidi all'informazione genetica che codifica per detti polipeptidi in un rgdp, detti rgdp esprimenti sulla loro superficie dette catene polipeptidiche.

rgdp è preferibilmente un fagemide esprimente uno o più polipeptidi legati a proteine di superficie.

I polipeptidi preparati con il processo secondo l'invenzione possono essere elementi di sbp e preferibilmente anticorpi specifici o loro frammenti. Preferibilmente, detti frammenti di anticorpi sono specifici scFv. I scFv, preferibilmente, avranno le regioni V collegate da un linker polipeptidico codificato almeno in parte dal sito loxP derivato dal fago P1, o un da un sito mutante loxP 511, derivato da loxP. Un esempio di linker, è il linker loxP 511 riportato in Figura 3 ed un esempio di polylinker è il pDAN5 polylinker riportato in Figura 4 il quale rappresenta una caratteristica dell'invenzione. Una ulteriore caratteristica dell'invenzione si riferisce a pDAN5 D1.3 descritto in Figura 5 e la cui



sequenza è riportata in Figura 6. Inoltre, l'invenzione si riferisce anche al polylinker pDAN5 di Figura 4 avente la sequenza nucleotidica corrispondente alla sequenza 1-2176 della sequenza riportata in Figura 6..

Una volta che le particelle infettanti o rdgp esprimono polipeptidi ricombinati sulla loro superficie, in modo che il genotipo ed il fenotipo siano accoppiati, esse possono essere selezionate e la diversità può essere stimata. In questo modo possono essere selezionati polipeptidi, ad esempio anticorpi o frammenti di anticorpi ricombinati, che risultano migliorati, rispetto a quelli conosciuti. Ad esempio, possono presentare una maggiore affinità di legame o esprimere una maggiore attività.

I polipeptidi ricombinanti ottenuti potranno quindi essere degli anticorpi o frammenti (ad esempio scFv) o potranno essere altri tipi di polipeptidi, come ad esempio quelli ottenuti, con metodi diversi, e descritti nello stato dell'arte.

Le regioni leganti dell'anticorpo si dividono naturalmente in catene VH e VL che possono venire unite fra di loro mediante un polipeptide codificato mediante un linker lox come descritto in questa sede. Tuttavia, questo metodo può anche venire usato per creare genoteche grandi di altra proteina (naturale o artificiale) presentata su fago (o meno) in cui può venire inserita una sequenza di ricombinazione sito-specifica (per esempio loxP o derivato), tale che possa venire creata una diversità nelle due parti della proteina separate dal linker, oppure in due catene polipeptidiche che possono venire unite covalentemente mediante tale linker. Inoltre, tali sequenze di ricombinazione sito-specifiche possono

anche venire inserite come sequenze non codificanti fra due catene polipeptidiche che vengono sintetizzate separatamente e diventano associate, covalentemente o meno, dopo la sintesi (per esempio catene Fab, recettori della cellula T, recettori della superficie cellulare).

Per la particolare applicazione a elementi di una coppia legante specifica che sono costituiti da una prima ed una seconda catena polipeptidica, che possono essere legate covalentemente o meno, sono necessarie due sequenze di ricombinazione sito-specifiche: una prima sequenza di ricombinazione sito-specifica ed una seconda sequenza di ricombinazione sito-specifica diversa dalla prima. La ricombinazione sito-specifica avvenendo fra le prime sequenze di ricombinazione sito-specifiche su molecole di acido nucleico differenti e fra le seconde sequenze di ricombinazione sito-specifiche su molecole di acido nucleico differenti, ma non fra una prima sequenza di ricombinazione sito-specifica ed una seconda sequenza di ricombinazione sito-specifica. Affinché possa verificarsi lo scambio di polipeptidi dell'elemento della coppia legante specifica fra le molecole di acido nucleico multiple introdotte in una singola cellula, lo scambio di geni della prima o seconda catena polipeptidica di elementi della coppia legante specifica, viene indotto nella prima sequenza di ricombinazione sito-specifica trovata fra il primo ed il secondo gene della catena polipeptidica ed una seconda sequenza di ricombinazione sito-specifica trovata fra il secondo ed il primo gene della catena polipeptidica (o viceversa). Il numero di sequenze di ricombinazione sito-specifiche che può venire usato, tuttavia, non è limitato a due: se ne possono anche usare tre o più,

provvedendo siti che soddisfano requisiti tali che possano ricombinarsi con se stessi e non con gli altri siti usati.

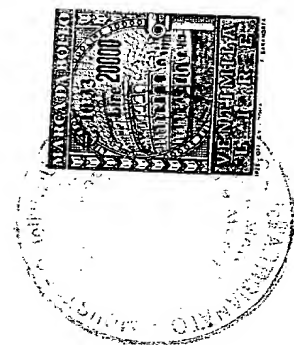
Sarà anche possibile selezionare polipeptidi leganti, ad esempio enzimi, anticorpi, ecc., facendoli legare a polipeptidi specifici, ad esempio facendoli passare su una colonna.

#### *Schema della Figura 2.*

Lo schema di ricombinazione con un solo (tipo di) vettore secondo la presente invenzione è illustrato schematicamente in Figura 2 e qui di seguito discusso. La ricombinazione è mostrata in modo schematico e per semplicità limitata a solo due vettori, in questo caso fagemidici, ma chiaramente l'infezione di una cellula può avvenire con un numero molto alto di vettori.

I substrati per la ricombinazione sono clonati in un singolo vettore backbone, in modo da costituire un fagemide contenente le origini di replicazione ColE1 e Ff. Si ottengono due fagemidi portanti rispettivamente i geni per VHA/VLA e VHB/VLB. Con la ricombinazione si ottengono 4 diversi fagemidi: i due fagemidi di partenza (VHA/VLA e VHB/VLB) e i due nuovi fagemidi (VHA/VLB e VHB/VLA). Differentemente dal metodo dello stato dell'arte riportato in Figura 1 in cui solo il 25% dei fagi ottenuti dalla ricombinazione porta anticorpi funzionali, nello schema di Figura 2 tutti i fagi portano anticorpi funzionali.

Come sopra già detto, lo schema di Figura 2 è mostrato in maniera semplificata. Infatti, una singola cellula batterica può essere infettata da un numero molto elevato di vettori, nell'Esempio 3, infatti, una singola



RP

cellula viene infettata da 19 fagemidi. Ne deriva che la potenzialità ricombinatoriale è molto alta, fino a 361 (19x19) volte nel caso dell'Esempio 3, e molto più elevata dei metodi descritti nello stato dell'arte WO 93/19172 o in Griffiths et al, 1994.

Inoltre, è chiaro che il metodo secondo l'invenzione può essere realizzato usando diversi tipi di cellule, geni e vettori infettanti, e diversi tipi e numero di polipeptidi. Per esempio si potrebbero selezionare percorsi metabolici che coinvolgono molte diverse proteine.

Come riportato in precedenza, L'uso di questo sistema non è limitato a E. coli e fagemidi. Sarà adatta qualsiasi particella infettiva in grado di infettare qualsiasi cellula in modo che copie multiple di acido nucleico possano entrare nella cellula e diventare substrati per la ricombinazione, e a questo proposito una particella infettiva può venire considerata come una molecola di acido nucleico rivestita con 'molecole di rivestimento' che sono in grado di entrare nelle cellule con elevata efficienza in virtù delle interazioni specifiche fra le molecole di rivestimento ed un recettore sulla superficie cellulare. Tali particelle infettive possono essere naturali, che sono state progettate per contenere substrati di ricombinazione, oppure particelle infettive progettate che non contengono l'informazione genetica necessaria per replicarsi, ma sono in grado di entrare specificamente nelle cellule. Un esempio di tale particelle infettiva è un fagemide.

Si possono usare cellule eucariote, comprendenti lievito e cellule di mammifero, come pure le cellule procariote, sia gram negative che gram positive, che possono venire infettata da particelle infettive che

normalmente infettano tali cellule, oppure sono state progettate o selezionate per infettare tali cellule. Le cellule usate possono essere normali, mutate per indurre maggiori entità di ricombinazione, oppure progettate per esprimere ricombinasi specifiche in grado di agire su sequenze di acido nucleico specifiche presenti all'interno delle molecole di acido nucleico contenute entro le particelle infettive.

I substrati usati per la ricombinazione possono essere substrati naturali, come varianti naturali, ottenuti da una o più specie, di una particolare sequenza di DNA, che può essere funzionale o meno, oppure possono essere substrati creati in vitro mediante amplificazione PCR, come è stato descritto per i metodi in vitro di mutazione ricombinante.

La presente invenzione verrà ora descritta secondo particolari realizzazioni negli Esempi che seguono.

**Esempio 1: I batteri possono venire infettati da più di un fagemide**

Come prima fase verso lo sviluppo di un sistema di ricombinazione intracellulare, si valuta il numero di fagemidi differenti che possono infettare singoli batteri, usando cinque fagemidi differenti che esprimono differenti resistenze agli antibiotici (ampicillina)(Bluescript: Stratagene, La Jolla), canamicina (pMPM-K1), tetraciclina (pMPM-T1), cloramfenicolo (pBSL121) e gentamicina (pBSL141). I plasmidi pMPM sono da (Mayer, M.P. 1995. A new set of useful cloning and expression vectors derived from pBlueScript. *Gene* 163:41-6) e i plasmidi pBSL da (Alexeyev, M., Shokolenko, I. and Croughan, T. 1995. Improved antibiotic-resistance gene cassettes and omega elements for *Escherichia coli* vector construction and in vitro deletion/insertion mutagenesis. *Gene*

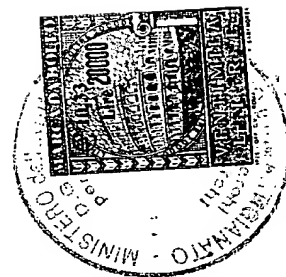
160:63-67). Infettando i batteri con titoli uguali di questi fagemidi e piastrandoli su antibiotici singoli o doppi, e considerando il numero di colonie presenti sulle piastre senza antibiotici come il 100%, si può vedere (Tabella 1) che la maggior parte dei batteri infettati con due fagemidi differenti ha due resistenze differenti.

**Tabella 1: ingresso e sopravvivenza di fagemidi, portanti resistenze agli antibiotici diverse, in una cellula**

Tetraciclina	100				
Canamicina	85	74			
Cloranfenicolo	70	63	64		
Gentamicina	73	33	39	52	
Ampicillina	94	81			92
	Tetraciclina	Canamicina	Cloranfenicolo	Gentamicina	Ampicillina

Per esempio, il 94% di tutti i batteri sono resistenti all'ampicillina ed alla tetraciclina, e l'85% sono resistenti alla tetraciclina ed alla canamicina. Come valutazione minima, il 10-40% dei batteri sembra avere il potenziale di presentare resistenza a tutti e cinque gli antibiotici. La resistenza alla gentamicina con altri antibiotici è stata trovata meno di frequente, ma questa è probabilmente una caratteristica di resistenza a questo antibiotico, dato che solo il 52% dei batteri può venire reso resistente alla gentamicina, anche quando usata da sola. La resistenza a certe coppie di antibiotici non viene indicata in quanto i plasmidi usati non ci permettono di valutare queste combinazioni.

Questi dati indicano che l'infezione, usando fagemidi, è un metodo



efficiente per introdurre almeno cinque diverse molecole di acido nucleico in una cellula.

**Esempio 2: La ricombinazione intracellulare può venire usata per costruire un anticorpo funzionale dalle sue catene componenti presenti su fagemidi separati**

Per usare la cre recombinasi per mescolare i geni delle regioni variabili nel formato scFv, occorre posizionare un sito lox fra i geni VH delle catene pesanti e VL delle leggere. Questo richiede l'uso del sito lox tradotto come linker di una proteina. Un esame delle sei cornici di lettura possibili disponibili per il sito loxP di tipo naturale e per il sito loxP 511 mutato (che non si ricombinerà con loxP di tipo naturale) (Hoess, R. and Abremski, K. 1985. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J Mol Biol* 181:351-362) identifica una traduzione di loxP 511 (ITSYNVYYTKL) che ha solo un singolo amminoacido basico (per ridurre la possibilità di proteolisi), manca dei codoni di arresto ed è il meno idrofobico. La sequenza del linker usato nel scFv è data nella Figura 3 e 4. La capacità di questa sequenza di agire come linker scFv nonché quella di altri due linker comunemente usati, (gly4ser)<sub>3</sub> (Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S. and Whitlow, M. 1988. Single-chain antigen-binding proteins [published erratum appears in *Science* 1989 Apr. 28;244(4903):409]. *Science* 242:423-426) e 220, una versione modificata del linker 218 (Whitlow et al., 1993), è stata ottenuta creando piccole genoteche scFv per ciascuno di questi linker e valutando i livelli

di esposizione mediante western blot sviluppando con l'anticorpo SV5 che individua il marcatore fra scFv clonato e la proteina del gene 3. Il linker loxP 511 mostra livelli visualizzati altrettanto buoni, se non migliori, degli altri due linker più ampiamente usati.

Sulla base di questo risultato, è stato progettato e costruito (Figure 4-6) un nuovo vettore d'esposizione fagemidico, pDAN5, e questo è stato provato per la capacità di esporre tre scFv differenti che derivano da anticorpi monoclonali precedentemente caratterizzati. I risultati indicano che tutti e tre i scFv individuano gli antigeni appropriati quando visualizzati in questo nuovo vettore usando il linker loxP 511. I segnali ELISA ottenuti usando questo vettore sono simili a quelli ottenuti usando vettori d'esposizione con il linker gly-ser.

Per ottenere pDAN5, un nuovo polylinker creato mediante PCR di sovrapposizione di due oligonucleotidi lunghi è stato clonato in pUC119 usando i siti di restrizione EcoRI e HindIII. Questo introduce una sequenza batterica leader, un sito multiplo di clonaggio contenente i siti di restrizione indicati in grassetto nella Figura 4, il marcatore SV5 (Hanke et al., 1992), un marcatore His<sub>6</sub> e un codone di stop amber (vedi Figure 4-6). Il gene 3 viene successivamente clonato in questo polylinker mediante amplificazione PCR da fdtet usando NotI e EcoRI. L'estremità 5' del gene 3 maturo viene inserita a valle del codone di arresto amber, e una sequenza lox di tipo naturale viene inserita, mediante PCR, all'estremità 3' del gene 3 dopo il codone di arresto e prima del sito EcoRI.

Il scFv D1.3 è stato costruito (nell'ordine VL-VH) da scFv D1.3 (ordine



VH-VL) amplificando VH D1.3 con VHback-DAN e VHfor-2-DAN, e VK D1.3 con VK2back-DAN e VK2for-DAN (tabella 2).

**Tabella 2: Primers usati per l'amplificazione PCR, oltre a quelli descritti in (Sblattero, D. and Bradbury, A. 1998. A definitive set of oligonucleotide primers for the amplification of human V regions, *Immunotechnology* 3:271-8)**

Oligo sequenza	Oligo nome
<u>TTA TCC TCG AGC GGT ACC</u> SAG GTS MAR CTG CAG SAG TCW GG	VHback-DAN
GAT TGG TTT GCC GCT AGC TGA GGA GAC GGT GAC CGT GGT	VHfor-2-DAN
<u>AGC AAG CGG CGC GCA TGC CGA</u> CAT CGA GCT CAC CCA GTC TC	VK2back-DAN
<u>GAA GTT ATG GTC GAC CCT CCG GAA</u> CGT TTG ATC TCG AGC TTG GTC	VK2forDAN
CC	
CGC TGG ATT GTT ATT ACT CGC AGC AAG CGG CGC GCA TGC C	VL back PT1
TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC	VL back PT2
CCA GGC CCA GCA GTG GGT TTG GGA TTG GTT TGC CGC TA	VH for PT1
TGG TGA TGG TGA GTA CTA TCC AGG CCC AGC AGT GGG TTT G	VH for PT2
ACC GCT CGA GGA TAA CTT CGT ATA GTA TAC ATT ATA CGA AGT TAT	VL for PTL
GGT CGA CCC TCC	
GGA GGG TCG ACC ATA ACT TCG TAT AAT GTA TAC TAT ACG AAG TTA	VH back PTL
TCC TCG AGC GGT A	

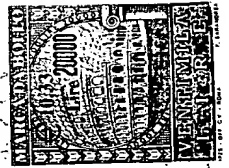
Tutti i geni V vengono purificati su gel e approssimativamente 200 ng vengono usati come stampi per una ulteriore amplificazione per aggiungere una regione di sovrapposizione nel linker scFv, come pure code lunghe per facilitare la digestione dell'enzima di restrizione. Gli oligonucleotidi VL back PT1 e VL for PTL vengono usati per amplificare i geni VL e VH for PT1 e VH back PTL per i geni VH. Le bande amplificate vengono purificate su gel come segue.

D1.3 scFv viene costruito miscelando quantità uguali (50-200 ng) dei geni VH e VL ed eseguendo la costruzione essenzialmente come

descritto in (Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., H.R., B. and Plückthun, A. 1997. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods*. 201:35-55): 8 cicli di pCR senza oligonucleotidi seguiti da 25 cicli in presenza di VL back PT2 e VH for PT2. I parametri del ciclo sono 94°C per 1 minuto (denaturazione), 60°C per 1 minuto (riassociazione) e 72°C per 1'30" (estensione).

Il scFv amplificato viene digerito con gli enzimi di restrizione BssHII e NheI e legato in pDAN5 tagliato con gli stessi enzimi. La miscela di legame viene elettroporata in DH5αF elettrocompetente e piastrata su piastre 2XTY 100 µg/ml di ampicillina/1% di glucosio, i cloni vengono confermati mediante sequenziamento. La capacità di mescolare intracellularmente i geni V a catena pesante e leggera per creare anticorpi funzionali viene valutata usando un scFv ottenuto dal mAb anti-lisozima, D1.3. Si creano, con clonazione PCR, due scFv che contengono D1.3 VH oppure D1.3 VL con catene partner irrilevanti, X e Y, (VL/D1.3-VH/X e VL/Y-VH/D1.3). Si è dimostrato che l'individuazione del lisozima mediante D1.3 scFv richiede la presenza delle catene pesanti e leggere di D1.3; singole catene di D1.3 associate con catene partner irrilevanti sono non funzionali. Un profilo dello schema usato è indicato nella Figura 7, secondo le fasi i)-ix):

i) 10 ml di E. coli BS1365, che esprime recombinasi cre costitutivamente (BS1365:BS591 F' kan (BS591:recA1 endA1 gyrA96 thi-1 D lacU169 supE44R17 [lambda1mm434 min5X1-cre]) (Sauer, B. and Henderson, N.



1988. The cyclization of linear DNA in *Escherichia coli* by site-specific recombination. *Gene* 70:331-341), vengono fatti crescere fino a DO550 0,5 a 37°C in 2XTY 100 µg/ml di canamicina/1% di glucosio.

ii) Quantità uguali di fagemide contenente i due geni scFv (VL/D1.3-VH/X e VL/Y-VH/D1.3) vengono aggiunte ai batteri ad un MOI di 20:1 ( $5 \times 10^{10}$  di ciascun fagemide aggiunto a  $5 \times 10^9$  batteri). Si lascia 30 minuti a 37°C senza agitazione per permettere che si verifichi l'infezione.

iii) Si aggiungono 100 µg/ml di ampicillina e i batteri vengono fatti crescere per almeno 12 ore a 30°C. La ricombinazione si verifica durante questo periodo.

iv) Dopo crescita per 12 ore, i batteri vengono diluiti 1:20 in 10 ml dello stesso mezzo di crescita, fatti crescere fino a DO550 0,5 e si aggiunge il fago helper M13 K07 a MOI di 20:1.

v) Si lascia per 30 minuti a 37°C senza agitazione per permettere che si verifichi l'infezione del helper fago.

vi) Si fa crescere la coltura per 6-18 ore a 30°C, si centrifuga a 4000 giri/minuto per 15 minuti e si raccoglie il supernatante. Questa fase prepara il fagemide che non espone scFv, o vi è un fenotipo non collegato al genotipo.

vii) Si fanno crescere 10 ml di DH5aF fino a DO 550 0,5 in 2XTY a 37°C.

viii) Il fagemide preparato nella fase 6 viene aggiunto a DH5aF a MOI inferiore a 1, lasciato per 30 minuti a 37°C e piastrato su piastre 2XTY 100 µg/ml di ampicillina/1% di glucosio. Questa fase accoppia fenotipo a genotipo, in assenza di questa fase l'scFv esposto può non necessariamente corrispondere al gene scFv presente nel fagemide.

ix) fagemidi esprimente scFv vengono prodotti da 20 colonie singole mediante crescita fino a DO550 0,5 in 10 ml di 2XTY 10 µg/ml di ampicillina/1% di glucosio, infezione con M13K07 a MOI 20:1 per 30 minuti a 37°C, centrifugazione e risospensione dei batteri in 10 ml di 2XTY 100 µg/ml di ampicillina e crescita per 6-18 ore a 30°C.

Se la ricombinazione avviene con successo, ciascun batterio dovrebbe contenere quattro geni scFv differenti (VL/D1.3-VH/D1.3; VL/D1.3-VH/X; VL/Y-VH/D1.3 e VL/Y-VHX) (vedi Figura 7). Si dimostra, mediante PCR, che il 25% (16/64) delle colonie batteriche ottenute nella fase viii/ix da cellule esprimenti la cre ricombinase presentano il corretto prodotto di ricombinazione. Inoltre, viene dimostrato che il 25% dei cloni fagemidici preparati nella fase ix riconoscono il lisozima in un test ELISA. Nei batteri che non esprimono cre, non è stata trovata ricombinazione e non è stato identificato alcun D1.3 funzionale, né per PCR né per ELISA. Questo dimostra che la ricombinazione viene indotta, come previsto, in quelle cellule che esprimono cre, tra ai siti loxP511 trovati fra VL e VH ed all'estremità del gene 3, il risultato netto risulta essere uno scambio di VH (e del gene 3 attaccato che è costante) fra fagemidi differenti.

**Esempio 3: Creazione di una grande genoteca di anticorpi in pDAN5 mediante la ricombinazione intracellulare.**

**Creazione di una genoteca scFv primaria mediante tecniche di clonazione standard**

Un genoteca primaria di fagemidi scFv costituita da  $7 \times 10^7$  cloni indipendenti, è stata creata in pDAN5 clonando scFv costruito con PCR ottenuto da geni Vµ, Vλ e Vκ di sangue periferico, nel modo seguente:

Linfociti di sangue periferico umano sono stati preparati mediante centrifugazione a gradiente di densità su Ficoll Hypaque (Pharmacia).

L'RNA totale viene preparato da questi linfociti mediante tiocianato di guanidinio acido, estrazione con fenolo/cloroformio e precipitazione con isopropanolo (Chomczynski, P. and Sacchi, N. 1987. Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159).

Si prepara cDNA usando trascrittasi inversa SuperScript II RNasi H<sup>-</sup> (Gibco BRL) con esameri casuali iniziando con 1-5 µg di RNA totale in un volume finale di 20 µl seguendo le istruzioni fornite con la SuperScript.

I geni IgM VH vengono dapprima amplificati con 0,5 µl di reazione con cDNA usando IgMfor e gli oligonucleotidi VHback descritti in (Sblattero e Bradbury, 1998). I volumi di reazione sono 20 µl, usando 0,5 µl di reazione con cDNA, 10 pmoli di ciascun oligonucleotide, dNTP 200 µM, tampone PCR 2 µl 10X e 0,5 µl (2,5 U) di Taq DNA polimerasi (Perkin Elmer). I parametri di reazione sono 94°C per 1 minuto (denaturazione), 55°C per 1 minuto (riassociazione) e 72°C per 1 minuto (estensione) per trenta cicli. Tutti i 20 µl vengono caricati su un gel di agarosio al 2% (FMC) e purificati su gel usando il kit di purificazione Qiagel (Qiagen). Successivamente, i geni VH vengono riamplicati usando miscele VHfor di oligonucleotidi con singoli oligonucleotidi VHback descritti in (Sblattero e Bradbury, 1998) in volumi di 50 µl usando 1 µl di VH purificato (gli altri parametri sono come in precedenza).

I geni Vλ e Vκ vengono amplificati in modo analogo (usando singoli

oligonucleotidi VLback con la miscela di oligonucleotidi VLfor) dal cDNA prodotto in precedenza.

Tutti i geni V vengono purificati su gel si approssimativamente 200 ng vengono usati come stampi per una ulteriore amplificazione per aggiungere una regione di sovrapposizione nel linker scFv, come pure code lunghe per facilitare la digestione dell'enzima di restrizione. Gli oligonucleotidi VL back PT1 e VL for PTL vengono usati per amplificare i geni VL e VH for PT1 e VH back PTL per i geni VH. Le bande amplificate vengono purificate da gel come in precedenza.

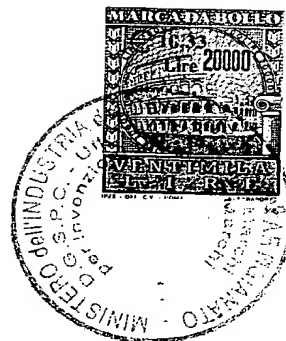
La genoteca scFv viene costruita miscelando quantità uguali (200-500 ng) di geni VH e VL ed eseguendo la costruzione essenzialmente come descritto in (Krebber et al., 1997): 8 cicli di PCR senza oligonucleotidi seguiti da 25 cicli in presenza di VL back PT2 e VH for PT2. I parametri dei cicli sono 94°C per 1 minuto (denaturazione), 60°C per 1 minuto (riassociazione) e 72°C per 1'30" (estensione).

I scFv amplificati vengono digeriti con BssHII e NheI e legato in pDAN5-D1.3 tagliato con BssHII/NheI. La miscela di legame viene elettroporata in DH5aF elettrocompetente e piastrata su piastre 2XTY 100 µg/ml di ampicillina/1% di glucosio per ottenere una genoteca primaria costituita da  $7 \times 10^7$  cloni indipendenti.

Le colonie vengono distaccate in 2XTY al 10% di glicerolo e congelate in aliquote di 1 ml.

#### **Ricombinazione intracellulare per creare una grande genoteca secondaria**

Per creare una grande genoteca secondaria viene seguito lo schema



illustrato nella Figura 8. Il protocollo dettagliato è il seguente:

Si diluiscono 100  $\mu$ l di batteri della genoteca primaria in 100 ml di 2XTY 100  $\mu$ g/ml di ampicillina/1% di glucosio e si fanno crescere fino ad DO550 0,5 e si aggiunge il fago helper M13 K07 a MOI di 20:1.

Il tutto viene lasciato per 30 minuti a 37°C senza agitazione per permettere che si verifichi l'infezione.

Si aggiunge canamicina a 100  $\mu$ g/ml e la coltura viene fatta crescere per 5-18 ore a 30°C, centrifugata a 4000 giri/minuto per 15 minuti e il supernatante raccolto. Questa fase prepara fagemidi che non espongono scFv, ma contengono i geni scFv.

20 ml di E. coli BS1365, che esprime costitutivamente ricombinasi cre, vengono fatti crescere fino a DO550 0,5 a 37°C in 2XTY 100  $\mu$ g/ml di canamicina/1% di glucosio.

Si aggiunge il fagemide preparato nella fase 3 a MOI di 100:1 ( $2 \times 10^{12}$  fagemidi vengono aggiunti a  $1 \times 10^{10}$  batteri). Si lascia per 30 minuti a 37°C senza agitazione per permettere che si verifichi l'infezione. Questo porta ad una infezione dei batteri da parte di più di un fagemide.

Si aggiunge ampicillina a 100  $\mu$ g/ml ed i batteri vengono fatti crescere per almeno 12 ore a 30°C. La ricombinazione si verifica durante questo periodo.

Dopo la crescita per almeno 12 ore, i batteri vengono aggiunti a 380 ml dello stesso mezzo di crescita (diluizione 1/20), fatti crescere fino a DO550 0,5 a 37°C e il fago helper M13 K07 viene aggiunto a MOI di 20:1. Dopo 30 minuti a 37°C senza agitazione, si aggiunge canamicina, i batteri vengono fatti crescere per altre 6-18 ore e il fagemide preparato

come descritto in precedenza. Questo produce una genoteca di fagemide con anticorpi ricombinati geneticamente diversi, ma senza proteine (che possono in ogni modo avere fenotipo e genotipo non accoppiati). Questo viene superato reinfettando DH5aF a MOI inferiore a 1.

1 l di DH5aF' viene fatto crescere fino a DO550 0,5 in 2XTY/1% di glucosio a 37°C, si aggiungono  $5 \times 10^{11}$  fagemidi (MOI inferiore a 1) e si lascia a riposo per 30 minuti per permettere che si verifichi l'infezione. Si aggiunge l'helper fago, M13K07, a MOI di 20:1, si lascia la coltura a riposo per 30 minuti a 37°C per permettere che si verifichi l'infezione e quindi si centrifuga a 4000 giri/minuto per 15 minuti. Si risospendono i batteri in 1 l di 2XTY 100 µg/ml di ampicillina/100 µg/ml di canamicina e si fanno crescere per 6-18 ore a 30°C. La coltura viene centrifugata (4000 giri/minuto, 30 minuti) ed il fagemide raccolto dal supernatante.

Si aggiungono al supernatante 150 ml di NaCl 2,5 M e PEG 8000 al 40% (questo precipita il fagemide). Il pellet viene recuperato mediante centrifugazione (4000 giri/minuto, 15 minuti), risospeso in 50 ml di PBS e centrifugato (9000 giri/minuto, 15 minuti) per rimuovere i batteri ed i fagemidi aggregati.

Si ripete la fase precedente aggiungendo 10 ml di NaCl 2,5 M e PEG 8000 al 40% al 50 ml di preparazione di fagemide. Il precipitato di PEG finale viene risospeso in 20 ml e il fagemide viene ulteriormente purificato mediante centrifugazione a densità di cloruro di cesio come descritto in (Smith, G.P. and Scott, J.K. 1993. Libraries of peptides and proteins displayed on filamentous phage. *Methods Enzymol* 217:228-57)



e risospeso in 20 ml. Questo costituisce la libreria anticorpale fagemidica finale che viene usata per le selezioni.

### **Valutazione di diversità in cellule singole**

Gli esperimenti descritti negli Esempi 1 e 2 precedenti indicano che almeno cinque fagemidi differenti possono entrare in singoli batteri e che due fagemidi che entrano in un singolo batterio sono in grado di ricombinarsi l'uno con l'altro fino all'equilibrio, ottenendo quattro fagemidi differenti. Tuttavia, essi non forniscono una indicazione di quanto a lungo tale fagemide possa sopravvivere dopo la crescita in coltura liquida o solida. Per chiarire questo punto, e caratterizzare la potenziale diversità della genoteca prodotta, fagemidi sono prodotti da una unica colonia di batteri cre contenente i fagemidi sottoposti alla ricombinazione, e questi fagemidi sono usati per ottenere colonie singole di batteri contenenti un unico tipo di fagemide. L'analisi di queste diverse colonie tramite PCR permette l'identificazione delle varie combinazioni di catene VH e VL presenti nella colonia cre di partenza (per i dettagli vedi Figura 8).

Dopo almeno 12 ore di crescita in batteri che esprimono cre, la coltura è stata piastrata per isolare le singole colonie. Se la ricombinazione ha avuto successo, e tutti i fagemidi sono ancora presenti, ciascuna di queste colonie dovrebbe contenere molti diversi fagemidi prodotti dalla ricombinazione. Per identificarli, è necessario preparare innanzitutto fagemidi da una singola colonia (questi dovrebbero essere costituiti da tutte le combinazioni scFv differenti che sono cresciute entro la cellula di partenza originale), e quindi usarle per preparare colonie contenenti

fagemidi singoli. In questo modo è possibile isolare i molti geni scFv singoli che si trovano nella colonia di partenza.

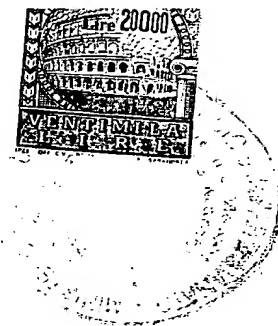
Dopo infezione a MOI elevato e crescita per una notte, i batteri che esprimono cre vengono piastrati su piastre 2XTY 100 µg/ml di ampicillina/1% di glucosio per isolare colonie singole.

I fagemidi vengono preparati da queste singole colonie facendo crescere tali colonie fino a DO550 0,5 in 1 ml di 2XTY 100 µg/ml di canamicina/100 µg/ml di ampicillina/1% di glucosio a 37°C, infettando con M13K07 e usando le tecniche suddescritte per produrre i fagemidi.

Per ottenere colonie da questi fagemidi, si fanno crescere batteri DH5aF fino a DO550 0,5 in 2XTY a 37°C e si infettano con i fagemidi ottenuti (nel paragrafo precedente) a MOI inferiore a 1. Tali colonie contengono singoli fagemidi e la popolazione completa di colonie rappresenta la diversità dei fagemidi contenuti entro i singoli batteri cre di partenza.

Le catene VH e VL singole presenti in ciascun fagemide vengono identificate mediante amplificazione PCR e mappatura con BstNI.

Due cellule vengono raccolte per l'analisi ed i risultati per entrambi sono mostrati nella Figura 9. Entrambe le cellule danno risultati molto simili, con 17-18 geni identificati per VH e 12-14 geni per VL. Questi geni V differenti sono presenti in 28-29 combinazioni differenti, con molti dei geni V differenti che si trovano in più di una combinazione scFv. Nel caso della prima cellula illustrata, si possono identificare tutte e quattro le combinazioni di due coppie VH/VL, indicando che la ricombinazione è molto estesa. In entrambe le cellule, oltre il 50% delle coppie VH/VL identificate sono presenti in coppie singole, suggerendo che il piccolo



campione analizzato (37-41 scFv) non ha identificato il complemento completo dei geni VH e VL e le loro combinazioni, e che la diversità vera è maggiore di quella identificata.

Il grado di diversità creato in una singola cellula è difficile da valutare. È molto probabile che i 18 geni VH differenti identificati in una singola colonia vengono accoppiati con 18 geni VL differenti. La potenziale diversità identificata in questo campione piccolo è 324 ( $18^2$ ). È improbabile che tutti i geni V differenti presenti vengano identificati, ed è noto che i geni V possono avere mappatura con BstN1 identiche, ma sequenze differenti, suggerendo che la diversità creata in una singola cellula può eccedere questo, ottenendo un massimo stimato che si avvicina al numero di 500-700 copie di pUC presenti in un singolo batterio (Sambrook et al., 1989). La valutazione della diversità più bassa è quella che è stata osservata e descritta nella figura 9, sulla base del numero di mappatura diversi: 28-29 differenti combinazioni scFv per cellula.

Il grado di ricombinazione è esteso, con molti gruppi differenti di tre combinazioni su quattro identificate. Una catena leggera è stata trovata con nove catene pesanti differenti e una catena pesante con cinque catene leggere differenti. Il fatto che molti plasmidi così differenti possano coesistere stabilmente in una singola cellula è alquanto sorprendente, e inoltre va contro il dogma che afferma che solo un plasmide per gruppo di compatibilità è in grado di sopravvivere in una cellula batterica. Inoltre, è probabile che quando un gran numero di geni V è stato introdotto in una cellula, le ricombinazioni che si verificano

mantengano la diversità e impediscono anche una eventuale pressione selettiva che tende ad eliminare singoli geni V. Questo fornisce un campo potenziale di diversità, creata mediante una singola cellula, di 29-500 scFv differenti. Nella genoteca creata in questa sede,  $10^{10}$  (20 ml) batteri che esprimono recombinasi cre vengono infettati con  $2 \times 10^{12}$  fagi (MOI di 200). Sulla base della diversità realmente osservata (ciascun batterio produce 29 fagi differenti), la dimensione della genoteca finale sarà  $2,9 \times 10^{11}$  ( $29 \times 10^{10}$ ). Se sono presenti tutte le 324 combinazioni possibili del gene V la genoteca sarà approssimativamente 10 volte più grande ( $3,2 \times 10^{12}$ ). Tuttavia, la dimensione della genoteca che può venire usata in pratica, viene limitata dalla fase di reinfezione a MOI basso quando fenotipo e genotipo sono accoppiati. Usando un litro per eseguire questa fase, non può eccedere  $5 \times 10^{11}$ , il numero di batteri in un litro.

#### **Test della grande genoteca fagica esprimente anticorpi**

La genoteca viene valutata mediante selezione su un certo numero di antigeni differenti (vedi Tabella 3) e anticorpi vengono isolati contro tutti in due cicli. Si ottengono in ogni caso almeno 3 scFv differenti per antigene con una media di 6 per antigene.

**Tabella 3: Antigeni per i quali sono stati selezionati anticorpi**

Antigene	% positive dopo 2-3 turni di selezione	numeri di scFv indipendenti	Affinità (nM)
<b>Umano</b>			
siero albumina	96	7	
Cyclin D	50	4	
Cdk2	96	9	29, 59, 82
Rad52	80	11	14
Flap endonucleasi	50		
Ku70/80	25		
cdc25A	25	3	
cdc25C	13		
PARP DNA dominio di legame	38		
PARP 85kDa	68		
<b>Rhodococcus</b>			
pigs10:ferrodossina	38		
piruvato ossireduttasi			
pigs12B: fosfoglicerato deidrogenasi	50	4	
<b>miscellanea</b>			
HIV1 loop	75	6	

La selezione viene eseguita come segue:

Si riveste una provetta (Nunc) con 1 ml dell'antigene di interesse a 10 µg/ml per una notte a 4°C. Dopo il rivestimento, si lava tre volte la provetta con PBS e si blocca riempiendo la provetta con Marvel 2%/PBS (MPBS). Simultaneamente, 1 ml di genoteca (contenente  $5 \times 10^{12}$ - $10^{13}$  fagemidi) viene bloccato aggiungendo Marvel 10%/PBS fino ad una concentrazione finale del 2%.

Dopo bloccaggio, la provetta viene lavata 3 volte con PBS e si aggiunge la genoteca, quindi si incuba per un'ora a riposo e un'ora con lenta agitazione a capovolgimento.

Dopo selezione, la provetta viene lavata dieci volte con PBS e dieci volte con Tween 20 0,15%/PBS.

Il fagemide di legame viene eluito con 1 ml di trietanolamina 100 mM per 8 minuti a temperatura ambiente e immediatamente trasferito in una nuova provetta e neutralizzato con Tris 1 M, pH 7-7,5.

Il fagemide eluito viene aggiunto a DH5aF a DO550 0,5 e lasciato senza agitazione per 30 minuti a 37°C.

I batteri infettati vengono piastrati su piastre 2XTY/100 µg/ml di ampicillina/1% di glucosio e fatti crescere per 12-18 ore a 30°C.

I batteri vengono raccolti in 10 ml di 2XTY, accuratamente miscelati e diluiti in 50 ml di 2XTY/100 µg/ml di ampicillina/1% di glucosio a DO550 0,05 e il fagemide viene preparato mediante crescita a 37°C fino a DO550 0,5 infettando con M13K07 e seguendo il protocollo descritto in precedenza, comprendente due precipitazioni con PEG/NaCl ottenendo un volume finale di 1 ml di fagemide.



$10^{12}$  di questi fagemidi vengono usati per la fase successiva di selezione che viene eseguita esattamente come descritto in precedenza, ad eccezione del fatto che nella fase di lavaggio il lavaggio viene eseguito venti volte con PBS e dieci volte con Tween 20 0,1%/PBS seguito da un lungo lavaggio di 30-60 minuti in PBS a temperatura ambiente.

I fagi eluiti vengono piastrati a bassa densità per isolare le singole colonie, il fagemide viene ottenuto da queste e valutato in ELISA come descritto in (Marks et al., 1991).

**Esempio 4: ricombinazione casuale in vivo non mediata da ricombinasi specifiche.**

L'enzima beta lattamasi non e' in grado di dare resistenza all'antibiotico cetaxina, tranne che ad una concentrazione di 0.025 mg/ml, concentrazione che non e' maggiore di quella in cui crescono ugualmente i batteri che non portano questo gene per la resistenza. E' stato dimostrato che la resistenza alla cefotaxime puo' essere ottenuta tramite mutazione del gene della beta lattamasi in alcuni residui chiave [Stemmer, 1994 #1870]. Tale dimostrazione e' avvenuta mediante ricombinazione in vitro e non e' stata invece ottenuta utilizzando solamente la tecnica della PCR incline all'errore. Per dimostrare una ricombinazione casuale in vivo non mediata da ricombinasi specifiche il gene della beta lattamasi (codificante per la resistenza all'ampicillina) e' stato amplificato dal vettore PBL5 167 [Alexeyev, 1995 #2287] utilizzando condizioni di PCR incline all'errore. Dopo trasformazione è stata ottenuta una libreria primaria di  $3 \times 10^5$  cloni. I fagemidi ottenuti da queste colonie sono stati utilizzati per infettare una popolazione di batteri

con diversi rapporti di fagi:batteri. I batteri infettati sono stati quindi cresciuti per almeno 12 ore ed in seguito da questi sono stati preparati nuovamente i fagemidi. E' stato trovato che i fagemidi preparati da i batteri infettati con un alto rapporto fagi:batteri (maggiore di 10:1) erano in grado di dare un numero elevato di colonie con resistenza ad una concentrazione di 0.25 mg/ml di cefotaxime, mentre i fagemidi preparati da batteri infettati con un basso rapporto fagi:batteri (minore di 1:1) erano in grado di dare molte meno colonie resistenti alla concentrazione di 0.25 mg/ml di cefotaxime. La ricombinazione puo' solamente avvenire quando più di un fagemide è presente all'interno di un singolo batterio, e una ricombinazione favorevole può avvenire solamente se almeno due fagemidi con mutazioni complementari (mutazioni che si trovano nello stesso gene della beta lattamasi, piuttosto che in due geni diversi, sono in grado di dar luogo ad una maggiore resistenza alla cefotaxime) sono presenti nello stesso batterio. Poichè l'unica differenza tra quei fagemidi che sono in grado di dare origine ad un numero maggiore di colonie resistenti rispetto a quelli in grado di dare un numero minore è il diverso rapporto di infezione fagi:batteri risulta evidente che la ricombinazione è avvenuta nel primo caso ma non nel secondo.

L'esatto protocolle utilizzato e' il seguente:

- 1 Il gene della beta lattamasi è stato amplificato utilizzando condizioni di PCR incline all'errore (REF) utilizzando due oligonucleotidi che ibridano nel polilinker del vettoreb PBL5 167
- 2 Il gene amplificato e' stato quindi tagliato con gli enzimi di restrizione Sac I e Kpn I ed e' stato riclonato nel vettore PBL5 167



tagliato con gli stessi enzimi, in modo da ottenere una libreria primaria di  $3 \times 10^5$  colonie in piastre cloramfenicolo (34 mg/ml). La resistenza al cloramfenicolo e' data dal gene cloramfenicolo acetiltransferasi che si trova nella parte costante (backbone) del vettore.

3 I batteri della libreria sono stati recuperati dalle piastre e cumulati, in seguito una parte è stata diluita in 10 ml di 2XTY ad una densità ottica di 0.1 a 550 nm. Dopo una crescita a 37°C una volta raggiunta una densità ottica di 0.5 550 nm sono stati aggiunti  $10^{11}$  fagi helper e dopo 30 minuti di incubazione a 37°C viene aggiunta la canamicina fino ad una concentrazione finale di 25 mg/ml .

4 I batteri vengono cresciuti a 30 °C per almeno 12 ore e i fagemidi sono recuperati dalla coltura. La coltura viene centrifugata (3500 giri al minuto per 20 minuti) e il sopranatante viene mescolato con 2.5 ml di NaCl 2.5 M e 40% Peg 6000 (questo precipita i fagemidi). Dopo centrifugazione (3000 giri al minuto per 20 minuti) il fondello ottenuto viene risospeso in 1 ml di PBS e centrifugato nuovamente (13000 giri al minuto per 10 minuti) per rimuovere residui batterici e fagemidi aggregati. Il sopranatante costituisce la libreria fagemidica primaria.

5 Diverse aliquote di 10 ml di batteri DH5aF' vengono cresciute fino ad una densità ottica a 550nm di 0.5 ed in seguito mescolate con i fagemidi della libreria primaria con diversi rapporti fagi:batteri varianti tra 0.1:1 fino a 100:1. I batteri sono lasciati per 30 minuti a 37 °C per permettere l'infezione e di seguito viene aggiunto il cloramfenicolo e i batteri sono cresciuti per almeno 12 ore a 30 °C.

6 Dopo la crescita i fagemidi sono preparati dalla coltura come

descritto nei punti 3-4. Tali fagemidi costituiscono la libreria secondaria o ricombinata.

7 Per verificare la capacità della libreria secondaria o ricombinata di dar luogo a batteri con una maggiore resistenza alla cefotaxime, batteri DH5aF' sono cresciuti fino ad una densità ottica a 550 nm di 0.5 e a diverse aliquote di 1 ml di questi vengono aggiunti i fagemidi preparati dalle diverse librerie secondarie (prodotte infettando con diversi rapporti fagi:batteri). I batteri vengono lasciati 30 minuti a 37 °C per permettere l'infezione e in seguito piastrati in piastre con diversa concentrazione di cefotaxime.

8 Librerie secondarie create da batteri infettati con rapporti fagi:batteri maggiori di 10:1 danno origine a circa 100-800 colonie su 0.25 mg/ml di cefotaxime, mentre sia la libreria primaria che le librerie create da batteri infettati con un rapporto fagi:batteri minore di 1 non danno luogo ad alcuna colonia resistente.

## RIVENDICAZIONI

1. Processo per la preparazione di una genoteca comprendente le seguenti fasi:

a) introduzione di almeno due membri di una popolazione di molecole di acido nucleico all'interno di almeno una cellula di una popolazione di cellule, detta popolazione di molecole di acido nucleico comprendente molecole individuali di acido nucleico ognuna delle quali essendo costituita da due parti: I) una sequenza di acido nucleico identica per ogni molecola; e II) una sequenza di acido nucleico variabile, non necessariamente contigua, la quale è il substrato per la ricombinazione;

b) ricombinazione avviene tra le molecole di acido nucleico inserite intracellularmente durante la fase a)

2. Processo secondo la rivendicazione 1, in cui detta ricombinazione della fase b) avviene tramite i sistemi ricombinatoriali presenti nella cellula.

3. Processo secondo la rivendicazione 1, in cui la ricombinazione della fase b) avviene tra siti specifici.

4. Processo secondo la rivendicazione 3, in cui la ricombinazione tra diverse molecole di acido nucleico avviene in uno o più siti specifici che si trovano all'interno di dette molecole di acido nucleico ed è catalizzata da ricombinasi, specifiche per detti siti, espresse dalla cellula.

5. Processo secondo le rivendicazioni 3-4, in cui detta ricombinazione avviene tra molecole diverse di acido nucleico su siti specifici loxP o da sequenze derivate da siti loxP, detti siti loxP essendo presenti in



*Handwritten signature or initials.*

dette molecole di acido nucleico, e dove detta ricombinazione è catalizzata dalla cre ricombinasi espressa dalla cellula.

6. Processo secondo le rivendicazioni 3-5, in cui l'acido nucleico introdotto nelle cellule e substrato per la ricombinazione comprende:

- i) una prima sequenza di ricombinazione sito specifica; e
- j) una seconda sequenza di ricombinazione sito specifica diversa dalla prima;

ed in cui la ricombinazione sito specifica avviene tra le prime sequenze di ricombinazione sito specifiche su diverse molecole di acido nucleico e tra le seconde sequenze di ricombinazione sito specifiche sulle stesse molecole di acido nucleico in cui avviene la ricombinazione al primo sito, ma non tra una prima sequenza di ricombinazione sito specifica ed una seconda sequenza di ricombinazione sito specifica.

7. Processo secondo la rivendicazione 6, in cui la prima sequenza di ricombinazione sito specifica è loxP e la seconda sequenza di ricombinazione sito specifica è un mutante della sequenza loxP.

8. Processo secondo la rivendicazione 7, in cui il mutante del sito loxP è loxP 511.

9. Processo secondo le rivendicazioni 1-8, in cui la cellula è una cellula batterica.

10. Processo secondo la rivendicazione 9, in cui la cellula batterica è Escherichia coli.

11. Processo secondo le rivendicazioni 1-10, in cui l'acido nucleico, substrato per la ricombinazione, viene introdotto nelle cellule tramite

trasfezione.

12. Processo secondo le rivendicazioni 1-10, in cui l'acido nucleico, substrato per la ricombinazione, è compreso all'interno di particelle infettanti e viene introdotto nelle cellule tramite infezione con dette particelle infettanti.
13. Processo secondo la rivendicazione 12 in cui le particelle infettanti sono fagi.
14. Processo secondo la rivendicazione 13 in cui le particelle infettanti sono fagi filamentosi.
15. Processo secondo la rivendicazione 14 in cui le particelle infettanti sono fagi filamentosi della famiglia Ff.
16. Processo secondo la rivendicazione 12 in cui le particelle infettanti sono fagemidi contenente DNA fagemidico.
17. Processo secondo la rivendicazione 16 in cui le particelle infettanti sono fagemidi derivati da fagi filamentosi della famiglia Ff contenente DNA fagemidico.
18. Processo secondo le rivendicazioni 1-17, in cui la parte con sequenza di acido nucleico variabile la quale è il substrato per la ricombinazione comprende sequenze per la regolazione genica.
19. Processo secondo le rivendicazioni 1-17, in cui la parte con sequenza di acido nucleico variabile la quale è il substrato per la ricombinazione comprende sequenze codificanti uno o più polipeptidi.
20. Processo secondo la rivendicazione 19, in cui detta sequenza variabile comprende geni codificanti una prima catena polipeptidica ed una seconda catena polipeptidica di un membro di una coppia di

elementi aventi tra di loro legami specifici (sbp) tale che con la ricombinazione si ottengono molecole di acidi nucleici diverse codificanti per nuove combinazioni del primo e del secondo polipeptide.

21. Processo secondo la rivendicazione 20, in cui la ricombinazione tra geni di detta prima e seconda catena di polipeptidi di sbp è indotta a livello di una prima sequenza di ricombinazione sito specifica che si trova tra i geni di dette prima e seconda catene polipeptidiche e a livello di una seconda sequenza di ricombinazione sito specifica che si trova tra i geni di dette seconda e prima catene polipeptidiche, o viceversa.
22. Processo per la preparazione di polipeptidi utilizzando genoteche preparate secondo le rivendicazioni 1-21.
23. Processo per la preparazione di polipeptidi comprendente le seguenti fasi:
  - a) viene preparata una genoteca secondo le rivendicazioni 1-21, dove le cellule in cui è avvenuta la ricombinazione esprimono uno o più polipeptidi codificati dall'acido nucleico substrato in esse contenuto;
  - b) selezione o screening di uno o più polipeptidi espressi nella fase a) per ottenere uno o più acidi nucleici che codificano per il polipeptide o polipeptidi oggetto della selezione o screening.
24. Processo secondo la rivendicazione 23, dove l'oggetto della selezione o screening è un'attività biologica codificata da uno o più polipeptidi.



25. Processo secondo la rivendicazione 23, dove l'oggetto della selezione o screening è un'attività enzimatica codificata da uno o più polipeptidi
26. Processo secondo la rivendicazione 23, in cui nella fase a) detti polipeptidi, codificati da acidi nucleici substrato, sono espressi sulla superficie delle cellule infettate; e nella fase b) avviene una selezione o screening del polipeptide o dei polipeptidi espressi in modo da fornire una popolazione di acidi nucleici codificanti uno o più polipeptidi oggetto della selezione o screening.
27. Processo secondo la rivendicazione 23, in cui nella fase a) detti polipeptidi, codificati da acidi nucleici, sono espressi sulla superficie di "replicable genetic display package" (rgdp); e nella fase b) avviene una selezione o screening del polipeptide o dei polipeptidi espressi in modo da fornire una popolazione di acidi nucleici codificanti uno o più polipeptidi oggetto della selezione o screening.
28. Processo secondo le rivendicazioni 23-27, in cui la fase a) è ripetuta con una popolazione di acidi nucleici o particelle infettive ottenuti dalla fase b).
29. Processo secondo la rivendicazione 27, in cui detto rgdp è un fagemide esprimente uno o più polipeptidi legati a proteine di superficie.
30. Processo per la preparazione di polipeptidi secondo le rivendicazioni 22-29, in cui detta genoteca, comprendente una molteplicità di particelle infettanti o rgdp in cui il fenotipo e il genotipo non sono accoppiati, è utilizzata per infettare cellule con un rapporto

particelle infettanti o rgdp:cellule minore di 1 e viene prodotta una seconda genoteca comprendente una seconda molteplicità di particelle infettanti o rgdp aventi il fenotipo accoppiato al genotipo.

31. Processo per la preparazione di polipeptidi secondo le rivendicazioni 22-30, in cui detti polipeptidi sono elementi sbp.
32. Processo per la preparazione di polipeptidi secondo la rivendicazione 31, in cui detti elementi sbp sono anticorpi specifici o loro frammenti.
33. Processo secondo la rivendicazione 32, in cui detti frammenti di anticorpi sono specifici scFv.
34. Processo secondo la rivendicazione 33, in cui detti scFv hanno le regioni V collegate da un linker polipeptidico almeno una parte del quale è codificato dal sito loxP derivato dal fago P1, o un da un sito mutante loxP 511, derivato da loxP.
35. Una genoteca preparata secondo le rivendicazioni 1-21.
36. Una genoteca preparata secondo le rivendicazioni 1-21, comprendente rgdp esprimenti polipeptidi.
37. Una genoteca secondo la rivendicazione 36, in cui rgdp sono fagi o fagemidi.
38. Una genoteca secondo la rivendicazione 36, in cui detti polipeptidi sono scFv, in cui VH e VL uniti da un linker polipeptidico codificato da un sito loxP derivato dal fago P1, o un suo mutante loxP 511, derivato loxP.
38. Un polipeptide preparato secondo le rivendicazioni 1-32.
39. Cellula ospite comprendente una genoteca secondo le rivendicazioni



34-37 o uno o più polipeptidi secondo la rivendicazione 38.

40. Un vettore rgdp, pDAN5, avente la seguente sequenza nucleotidica:

AAGCTTGCCA AATTCTATTT CAAGGAGACA GTCATAATGA AATACCTATT GCCTACGGCA  
GCCGCTGGAT TGTATTACT CGCAGCAAGC GGC GCGCATG CCGACATTCA GATGACCCAG  
TCTCCAGCCT CCCTTTCTGC GTCTGTGGGA GAAACTGTCA CCATCACATG TCGAGCAAGT  
GGGAATATTC ACATTTATTT AGCATGGTAT CAGCAGAAAC AGGGAAAATC TCCTCAGCTC  
CTGGTCTATT ATACAACAAC CTTAGCAGAT GGTGTGCCAT CAAGGTTTCA TGGCAGTGGA  
TCAGGAACAC AATATTCTCT CAAGATCAAC AGCCTGCAAC CTGAAGATTT TGGGAGTTAT  
TACTGTCAAC ATTTTGGAG TACTCCTCGG ACGTTCGGTG GAGGGACCAA GCTGGAGCTG  
AAACGTTCCG GAGGGTCGAC CATAACTTCG TATAATGTAT ACTATACGAA GTTATCCTCG  
AGCGGTACCG AGGTGAAGCT GGTGGAGTCA GGACCTGGCC TGGTGGCGCC CTCACAGAGC  
CTGTCCATCA CATGCACCGT CTCAGGGTTC TCATTAACCG GCTATGGTGT AACTGGGTT  
CGCCAGCCTC CAGGAAAGGG TCTGGAGTGG CTGGGAATGA TTTGGGGTGA TGGAAACACA  
GACTATAATT CAGCTCTCAA ATCCAGACTG AGCATCAGCA AGGACAACTC CAAGAGCCAA  
GTTTCTTAA AAATGAACAG TCTGCACACT GATGACACAG CCGTCTACTA CTGCGCGCGA  
GAGAGAGATT ATAGGCTTGA CTACTGGGGC CAAGGCACCA CGGTCAACCGT CTCCTCAGCT  
AGCGGCAAAC CAATCCCAA CCCACTGCTG GGCCTGGATA GTACTCACCA TCACCATCAC  
CATTAGGCGG CCGCTACTGT TGAAAGTTGT TTAGCAAAAC CTCATACAGA AAATTCATTT  
ACTAACGTCT GGAAAGACGA CAAACTTTA GATCGTTACG CTAACATGA GGGCTGTCTG  
TGGAATGCTA CAGGCGTTGT GGTGTGTACT GGTGACGAAA CTCAGTGTTA CGGTACATGG  
GTTCTATTG GGCTTGCTAT CCCTGAAAAT GAGGGTGGTG GCTCTGAGGG TGGCGGTTCT  
GAGGGTGGCG GTTCTGAGGG TGGCGGTACT AAACCTCCTG AGTACGGTGA TACACCTATT  
CCGGGCTATA CTTATATCAA CCCTCTCGAC GGCACCTATC CGCCTGGTAC TGAGCAAAAC  
CCCCTAATC CTAATCCTTC TCTTGAGGAG TCTCAGCCTC TTAATACTTT CATGTTTCAG  
AATAATAGGT TCCGAAATAG GCAGGGTGCA TTAAGTGT TACGGGCAC TGTTACTCAA  
GGCACTGACC CCGTTAAAAC TTATTACCAG TACACTCCTG TATCATCAAA AGCCATGTAT  
GACGCTTACT GGAACGGTAA ATTCAGAGAC TGCGCTTTCC ATTCTGGCTT TAATGAGGAT  
CCATTCGTTT GTGAATATCA AGGCCAATCG TCTGACCTGC CTCAACCTCC TGTCAATGCT  
GGCGGCGGCT CTGGTGGTGG TTCTGGTGGC GGCTCTGAGG GTGGCGGCTC TGAGGGTGGC  
GGTTCTGAGG GTGGCGGCTC TGAGGGTGGC GGTTCGGTG GCGGCTCCGG TTCCGGTGAT  
TTTGATTATG AAAAAATGGC AAACGCTAAT AAGGGGGCTA TGACCGAAAA TGCCGATGAA  
AACGCGCTAC AGTCTGACGC TAAAGGCAAA CTGATTCTG TCGCTACTGA TTACGGTGCT

GCTATCGATG GTTTCATTGG TGACGTTTCC GGCCTTGCTA ATGGTAATGG TGCTACTGGT  
GATTTTGCTG GCTCTAATTC CCAAATGGCT CAAGTCGGTG ACGGTGATAA TTCACCTTTA  
ATGAATAATT TCCGTCAATA TTTACCTTCT TTGCCTCAGT CGGTTGAATG TCGCCCTTAT  
GTCTTTGGCG CTGGTAAACC ATATGAATTT TCTATTGATT GTGACAAAAT AAACCTTATC  
CGTGGTGTCT TTGCGTTTCT TTTATATGTT GCCACCTTTA TGTATGTATT TTCGACGTTT  
GCTAACATAC TGCCTAATAA GGAGTCTTAA GCATGCATAA CTTCTGTATAA TGTATGCTAT  
ACGAAGTTAT GAATTCACCTG GCCGTGCTTT TACAACGTCG TGACTGGGAA AACCTGGCG  
TTACCCAACT TAATCGCCTT GCAGCACATC CCCCTTTCGC CAGCTGGCGT AATAGCGAAG  
AGGCCCCGAC CGATCGCCCT TCCCAACAGT TCGCGAGCCT GAATGGCGAA TGGCGCCTGA  
TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTTACA CCGCATACGT CAAAGCAACC  
ATAGTACGCG CCCTGTAGCG GCGCATTAAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT  
GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCTTTTCT  
CGCCACGTTT GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTCCG  
ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TTGGGTGATG GTTCACGTAG  
TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA  
TAGTGGAATC TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCTCGGGCT ATTCTTTTGA  
TTTATAAGGG ATTTTGCCGA TTTGCGCTTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA  
ATTTAACGCG AATTTTAACA AAATATTAAC GTTTACAATT TTATGGTGCA CTCTCAGTAC  
AATCTGCTCT GATGCCGCAT AGTTAAGCCA GCGCCGACAC CCGCCAACAC CCGCTGACGC  
GCCCTGACGG GCTTGCTCTG TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG  
GAGCTGCATG TGTGAGAGGT TTTACCGTC ATCACCGAAA CGCGCGAGAC GAAAGGGCCT  
CGTGATACGC CTATTTTAT AGGTTAATGT CATGATAATA ATGGTTTCTT AGACGTCAGG  
TGGCACTTTT CGGGGAAATG TGCGCGGAAC CCCTATTTGT TTATTTTCTT AAATACATTC  
AAATATGTAT CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT ATTGAAAAAG  
GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT CCCTTTTTTG CGGCATTTTG  
CCTTCCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG AAGATCAGTT  
GGGTGCACGA GTGGGTTACA TCGAACTGGA TCTCAACAGC GGTAAAGATCC TTGAGAGTTT  
TCGCCCCGAA GAACGTTTTT CAATGATGAG CACTTTTAAA GTTCTGCTAT GTGGCGCGGT  
ATTATCCCGT ATTGACGCCG GGCAAGAGCA ACTCGGTCGC CGCATACACT ATTCTCAGAA  
TGACTTGTT GAGTACTCAC CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG  
AGAATTATGC AGTGCTGCCA TAACCATGAG TGATAACACT GCGGCCAACT TACTTCTGAC  
AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGAC AACATGGGGG ATCATGTAAC  
TCGCCTTGAT CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG AGCGTGACAC  
CACGATGCCT GTAGCAATGG CAACAACGTT GCGCAAACCTA TTAAGTGGCG AACTACTTAC

VDP



TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG GATAAAGTTG CAGGACCACT  
TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG CCGGTGAGCG  
TGGGTCTCGC GGTATCATTG CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT  
TATCTACACG ACGGGGAGTC AGGCAACTAT GGATGAACGA AATAGACAGA TCGCTGAGAT  
AGGTGCCTCA CTGATTAAGC ATTGGTAACT GTCAGACCAA GTTTACTCAT ATATACTTTA  
GATTGATTTA AAACCTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC TTTTGTATAA  
TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCCGTAGA  
AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC  
AAAAAAACCA CCGCTACCAG CCGTGGTTTG TTTGCCGAT CAAGAGCTAC CAACTCTTTT  
TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT ACTGTCCTTC TAGTGTAGCC  
GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG CTCTGCTAAT  
CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT TGGACTCAAG  
ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT GCACACAGCC  
CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC ATTGAGAAAG  
CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC  
AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG  
GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC TCGTCAGGGG GCGGAGCCT  
ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCTTG GCCTTTTGCT GGCCTTTTGC  
TCACATGTTT TTTCTGCGT TATCCCCTGA TTCTGTGGAT AACCGTATTA CCGCCTTTGA  
GTGAGCTGAT ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG TGAGCGAGGA  
AGCGGAAGAG CGCCAATAC GCAAACCGCC TCTCCCCGCG CGTTGGCCGA TTCATTAATG  
CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGGCAGT GAGCGCAACG CAATTAATGT  
GAGTTAGCTC ACTCATTAGG CACCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATGTT  
GTGTGGAATT GTGAGCGGAT AACAAATTCA CACAGGAAAC AGCTATGACC ATGATTACGC  
C.

41. Polylinker pDAN5 avente la sequenza nucleotidica 1-2176 riportata  
alla rivendicazione 40 e di sequenza:

AAGCTTGCCA AATTCTATTT CAAGGAGACA GTCATAATGA AATACCTATT GCCTACGGCA  
GCCGCTGGAT TGTTATTACT CGCAGCAAGC GGCGCGCATG CCGACATTCA GATGACCCAG  
TCTCCAGCCT CCCTTTCTGC GTCTGTGGGA GAAACTGTCA CCATCACATG TCGAGCAAGT  
GGGAATATTC ACATTTATTT AGCATGGTAT CAGCAGAAAC AGGGAAAATC TCCTCAGCTC  
CTGGTCTATT ATACAACAAC CTTAGCAGAT GGTGTGCCAT CAAGGTTTCA TGGCAGTGGA  
TCAGGAACAC AATATTCTCT CAAGATCAAC AGCCTGCAAC CTGAAGATTT TGGGAGTTAT

TACTGTCAAC ATTTTGGAG TACTCCTCGG ACGTTCGGTG GAGGGACCAA GCTGGAGCTG  
AAACGTTCCG GAGGGTCGAC CATAACTTCG TATAATGTAT ACTATACGAA GTTATCCTCG  
AGCGGTACCG AGGTGAAGCT GGTGGAGTCA GGACCTGGCC TGGTGGCGCC CTCACAGAGC  
CTGTCCATCA CATGCACCGT CTCAGGGTTC TCATTAACCG GCTATGGTGT AAAGTGGGTT  
CGCCAGCCTC CAGGAAAGGG TCTGGAGTGG CTGGGAATGA TTTGGGGTGA TGGAAACACA  
GACTATAATT CAGCTCTCAA ATCCAGACTG AGCATCAGCA AGGACAACTC CAAGAGCCAA  
GTTTTCTTAA AAATGAACAG TCTGCACACT GATGACACAG CCGTCTACTA CTGCGCGCGA  
GAGAGAGATT ATAGGCTTGA CTACTGGGGC CAAGGCACCA CGGTACCGT CTCCTCAGCT  
AGCGGCAAAC CAATCCCAA CCCACTGCTG GGCCTGGATA GTACTACCA TCACCATCAC  
CATTAGGCGG CCGCTACTGT TGAAAGTTGT TTAGCAAAAC CTCATACAGA AAATTCATTT  
ACTAACGTCT GGAAAGACGA CAAAACTTTA GATCGTTACG CTAAGTATGA GGGCTGTCTG  
TGGAATGCTA CAGGCGTTGT GGTGTTGACT GGTGACGAAA CTCAGTGTTA CGGTACATGG  
GTTCTTATG GGCTTGCTAT CCCTGAAAAT GAGGGTGGTG GCTCTGAGGG TGGCGGTTCT  
GAGGGTGGCG GTTCTGAGGG TGGCGGTACT AAACCTCCTG AGTACGGTGA TACACCTATT  
CCGGGCTATA CTTATATCAA CCCTCTCGAC GGCACCTATC CGCCTGGTAC TGAGCAAAAC  
CCCGCTAATC CTAATCCTTC TCTTGAGGAG TCTCAGCCTC TTAATACTTT CATGTTTCAG  
AATAATAGGT TCCGAAATAG GCAGGGTGCA TTAAGTGTGTT ATACGGGCAC TGTTACTCAA  
GGCACTGACC CCGTTAAAAC TTATTACCAG TACACTCCTG TATCATCAA AGCCATGTAT  
GACGCTTACT GGAACGGTAA ATTCAGAGAC TGCCTTTCC ATTCTGGCTT TAATGAGGAT  
CCATTGTTTT GTGAATATCA AGGCCAATCG TCTGACCTGC CTCAACCTCC TGTCAATGCT  
GGCGGCGGCT CTGGTGGTGG TTCTGGTGGC GGCTCTGAGG GTGGCGGCTC TGAGGGTGGC  
GGTTCTGAGG GTGGCGGCTC TGAGGGTGGC GGTTCGGTG GCGGCTCCGG TTCCGGTGAT  
TTTGATTATG AAAAAATGGC AAACGCTAAT AAGGGGGCTA TGACCGAAAA TGCCGATGAA  
AACGCGCTAC AGTCTGACGC TAAAGGCAA CTTGATTCTG TCGCTACTGA TTACGGTGCT  
GCTATCGATG GTTTCATTGG TGACGTTTCC GGCCTTGCTA ATGGTAATGG TGCTACTGGT  
GATTTTGCTG GCTCTAATTC CCAAATGGCT CAAGTCGGTG ACGGTGATAA TTCACCTTTA  
ATGAATAATT TCCGTCAATA TTTACCTTCT TTGCCTCAGT CGGTTGAATG TCGCCCTTAT  
GTCTTTGGCG CTGGTAAACC ATATGAATTT TCTATTGATT GTGACAAAAT AAAGTTATTC  
CGTGGTGTCT TTGCGTTTCT TTTATATGTT GCCACCTTTA TGTATGTATT TTCGACGTTT  
GCTAACATAC TCGTAATAA GGAGTCTTAA GCATGCATAA CTTCGTATAA TGTATGCTAT  
ACGAAGTTAT GAATTC


(MAT/pd)

12/11

Milano, li 18 Novembre 1998

p. S.I.S.S.A. - SCUOLA INTERNAZIONALE SUPERIORE DI  
STUDI AVANZATI

il Mandatario



Dr. Diego Pallini

NOTARBARTOLO & GERVASI S.p.A.

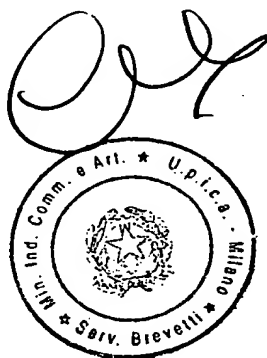
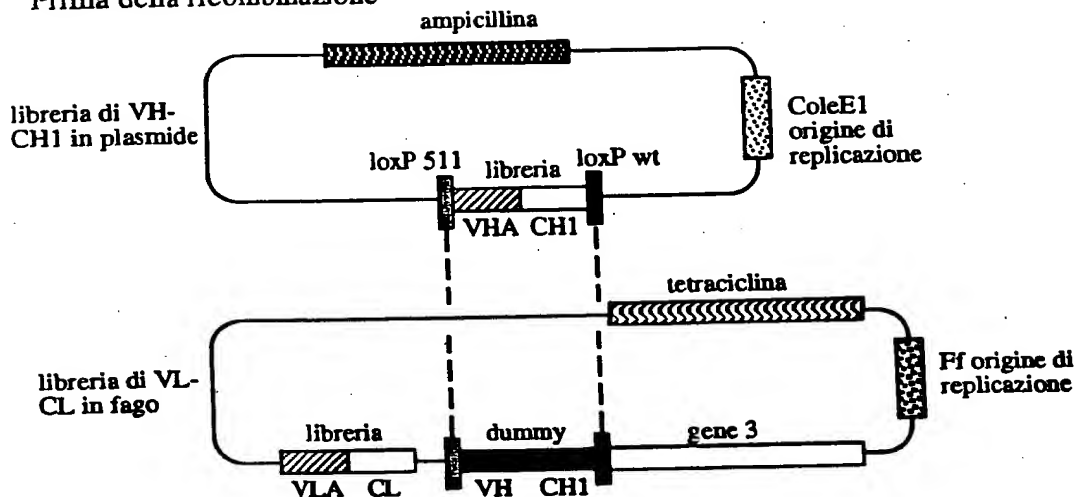
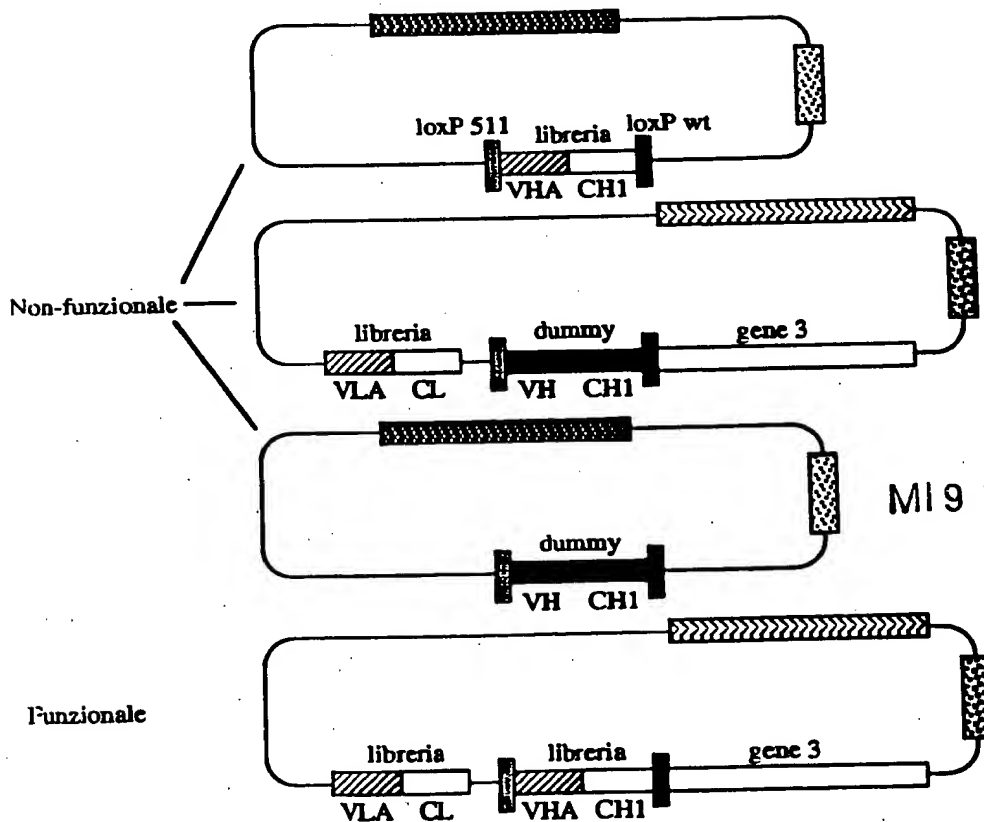


Figura 1

Prima della ricombinazione



Dopo ricombinazione



MI98A002509

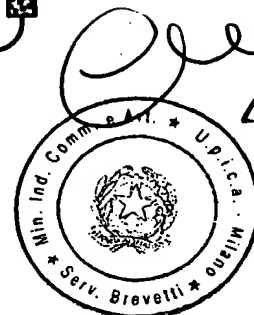
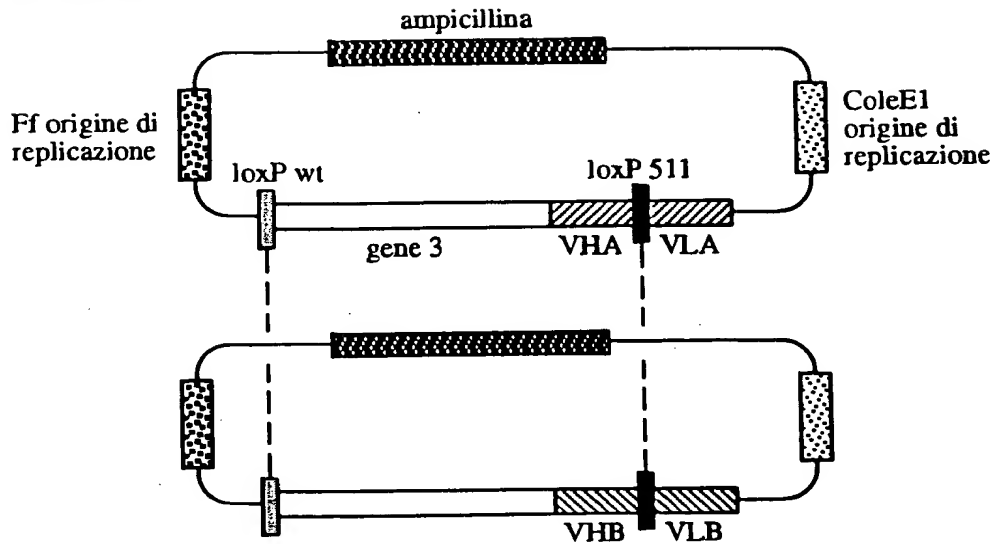


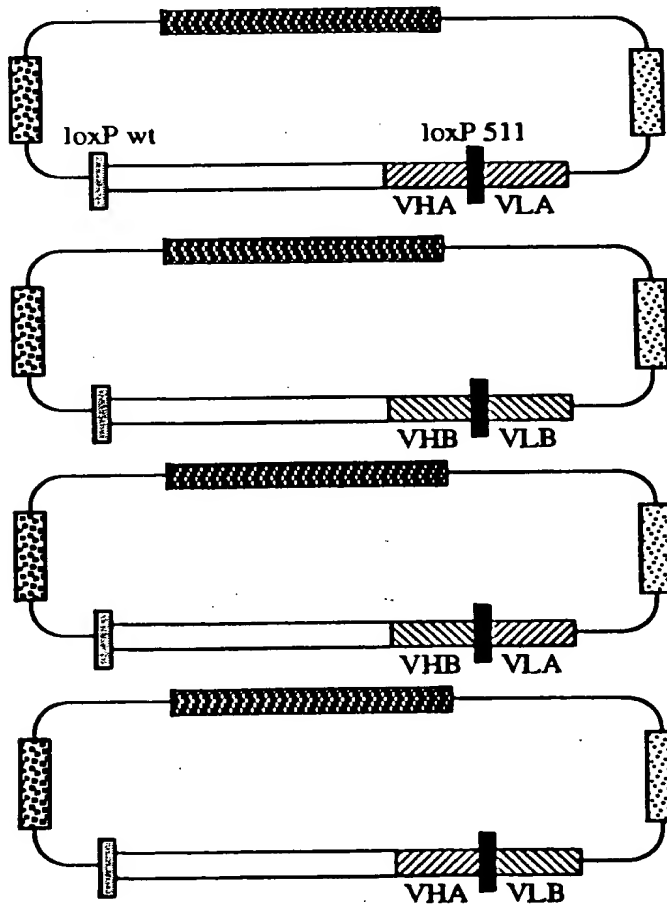


Figura 2

Prima della ricombinazione



Dopo ricombinazione



MI98A002509

Tutti i fagemidi, prima e dopo la ricombinazione, hanno scFv funzionali



1245

*[Handwritten signature]*

Figura 3

linker loxP 511

S G G S T I T S Y N V Y Y T K L S S S G T  
TCCGGAGGTCGACC ATAACTCGTATAATGTATACGAAGTTAT CCTCGAGCGGTACC  
 BspEI Sall Sito di ricombinazione loxP 511 XhoI KpnI

M198A002509

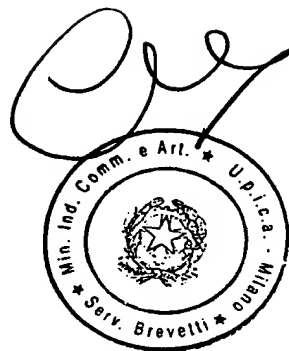




Figura 4

Polylinker del vettore pDAN5

leader batterico  
M K Y L L P T A A A G L L L A A S G A H A  
AAGCTTGGCCAAATTCATTTCAGGAGACAGTCATA ATGAATACCTATATGCTACGCGAGCGCTGGATTTATTACTCGCAGCAAGCGCGCGCATGCC  
HindIII Shine Delgano BspHI

S G G S T I T S Y N V Y T K L S S G T  
....catena VL...TCGGAGGGTGGACC ATAACTTGGTATATGCTACTATACGAAGTTAT CCTCGAGCGGTACC.....catena VH....  
BspEI SalI sito di ricombinazione loxP 511 XhoI KpnI

SV5 epitope tag His6  
A S G K P I P N P L L G L D S T H H H H H (Q)  
GCTAGC GGCAACCAATCCCAACCCACTGCTGGGCGCTGGAT AGTACT CACCATCACCATCACCAT TAG  
NheI ScaI amber

proteina del gene 3  
A A A T V E S C . . . R N K E S stop  
GGGCGCG TACGTGTAAGTTGT.....CGTAATAGGAGICT TAA GCATGC ATAACTTCGTATATGCTATGCTATACGAAGTTAT GAATTC  
NotI sito di ricombinazione loxP wt EcoRI

M198A002509

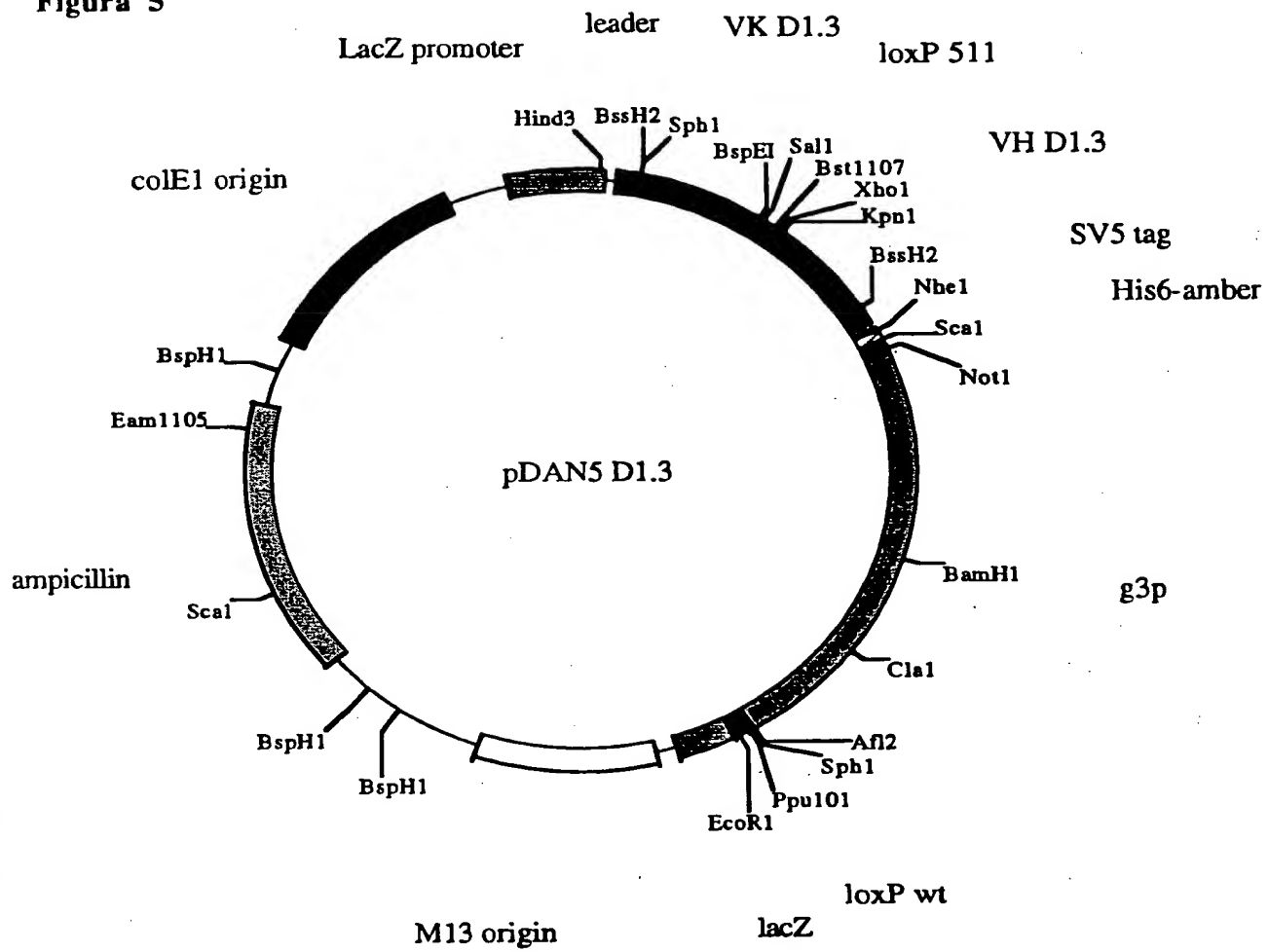


*[Handwritten signature]*

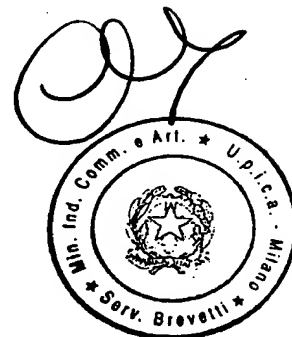
MAF

*F. Gerassi*

Figura 5



M19 8 A 002509



MAF

Figura 6a

## Sequenza del vettore pDAN5 D1.3

AAGCTTGCCA AATTCTATTT CAAGGAGACA GTCATAATGA AATACCTATT GCCTACGGCA  
 GCCGCTGGAT TGTATTACT CGCAGCAAGC GGCGCGCATG CCGACATTCA GATGACCCAG  
 TCTCCAGCCT CCCTTTCTGC GTCTGTGGGA GAAACTGTCA CCATCACATG TCGAGCAAGT  
 GGAATATTC ACATTTATTT AGCATGGTAT CAGCAGAAAC AGGGAAAATC TCCTCAGCTC  
 TGGGTCTATT ATACAACAAC CTTAGCAGAT GGTGTGCCAT CAAGGTTTCA TGGCAGTGGA  
 TCAGGAACAC AATATTCTCT CAAGATCAAC AGCCTGCAAC CTGAAGATTT TGGGAGTTAT  
 TACTGTCAAC ATTTTTGGAG TACTCCTCGG ACGTTCGGTG GAGGGACCAA GCTGGAGCTG  
 AAACGTTCCG GAGGGTCGAC CATAACTTCG TATAATGTAT ACTATACGAA GTTATCCTCG  
 AGCGGTACCG AGGTGAAGCT GGTGGAGTCA GGACCTGGCC TGGTGGCGCC CTCACAGAGC  
 CTGTCCATCA CATGCACCGT CTCAGGGTTC TCATTAACCG GCTATGGTGT AAACCTGGTT  
 CGCCAGCCTC CAGGAAAGGG TCTGGAGTGG CTGGGAATGA TTTGGGGTGA TGGAAACACA  
 GACTATAATT CAGCTCTCAA ATCCAGACTG AGCATCAGCA AGGACAACCT CAAGAGCCAA  
 GTTTTCTTAA AAATGAACAG TCTGCACACT GATGACACAG CCGTCTACTA CTGCGCGCGA  
 GAGAGAGATT ATAGGCTTGA CTACTGGGGC CAAGGCACCA CCGTCACCGT CTCCTCAGCT  
 AGCGGCAAAAC CAATCCCAAA CCCACTGCTG GGCCTGGATA GTACTACCA TCACCATCAC  
 CATTAGGCGG CCGCTACTGT TGAAAGTTGT TTAGCAAAAC CTCATACAGA AAATTCATTT  
 ACTAACGTCT GGAAAGACGA CAAAACCTTA GATCGTTACG CTAACATGA GGGCTGTCTG  
 TGGAAATGCTA CAGCGTGTGT GGTTTGTACT GGTGACGAAA CTCAGTGTTA CCGTACATGG  
 GTTCTTATTG GGCTTGCTAT CCCTGAAAAT GAGGGTGGTG GCTCTGAGGG TGGCGGTTCT  
 GAGGGTGGCG GTTCTGAGGG TGGCGGTACT AAACCTCCTG AGTACGGTGA TACACCTATT  
 CCGGGCTATA CTTATATCAA CCCTCTCGAC GGCACCTATC CGCCTGGTAC TGAGCAAAAC  
 CCCGCTAATC CTAATCCTTC TCTTGAGGAG TCTCAGCCTC TTAATACTTT CATGTTTCAG  
 AATAATAGGT TCCGAAATAG GCAGGGTGCA TTAACCTGTT ATACGGGCAC TGTTACTCAA  
 GGCACGTACC CCGTTAAAAC TTATTACCAG TACACTCCTG TATCATCAA AGCCATGTAT  
 GACGCTTACT GGAACGGTAA ATTACAGAGC TGCGCTTTCC ATTCTGGCTT TAATGAGGAT  
 CCATTGCTTT GTGAATATCA AGGCCAATCG TCTGACCTGC CTCAACCTCC TGTCATGTCT  
 GGCGGCGGCT CTGGTGGTGG TTCTGGTGGC GGCTCTGAGG GTGGCGGCTC TGAGGGTGGC  
 GGTCTGAGG GTGGCGGCTC TGAGGGTGGC GGTTCGGTG TACCGGAAAT TGCCGGTGAT  
 TTTGATTATG AAAAAATGGC AAACGCTAAT AAGGGGGCTA TGACCGAAAA TGCCGATGAA  
 AACGCGCTAC AGTCTGACGC TAAAGGTGCA CTGATTCTG TCGCTACTGA TTACGGTGCT  
 GCTATCGATG GTTTCATTGG TGACGTTTCC GGCCTTGCTA ATGGTAATGG TGCTACTGGT  
 GATTTTGCTG GCTCTAATTC CCAAATGGCT CAAGTCGGTG ACGGTGATAA TTCACCTTTA  
 ATGAATAATT TCCGTCAATA TTTACCTTCT TTGCCTCAGT CCGTTGAATG TCGCCCTTAT  
 GTCTTTGGCG CTGGTAAACC ATATGAATTT TCTATTGATT GTGACAAAAT AAACCTATT  
 CGTGGTGTCT TTGCGTTTCT TTTATATGTT GCCACCTTTA TGTATGTATT TTCGACGTTT  
 GCTAACATAC TGCGTAATAA GGAGTCTTAA GCATGCATAA CTTCGTATAA TGTATGCTAT  
 ACGAAGTTAT GAATTCACCT GCCGTCGTTT TACAACGTCG TGACTGGGAA AACCTGGCG  
 TTACCCAACT TAATCGCCTT GCAGCACATC CCCCTTTTCG CAGCTGGCGT AATAGCGAAG  
 AGGCCCGCAC CGATCGCCCT TCCCAACAGT TGCGCAGCCT GAATGGCGAA TGGCGCCTGA  
 TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTTACA CCGCATACGT CAAAGCAACC  
 ATAGTACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT  
 GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT  
 CGCCACGTTT GCCGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG  
 ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TTGGGTGATG GTTCACGTAG  
 TGGGCCATCG CCCTGATAGA CCGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA  
 TAGTGGACTC TTGTTCCTAA CTGGAACAAC ACTCAACCCT ATCTCGGGCT ATTCCTTTGA  
 TTTATAAGGG ATTTTGCCGA TTTCCGCCCTA TTGGTTAAAA AATGAGCTGA TTTAACAATA  
 ATTTAACGCG AATTTTAACA AAATATTAAC GTTTACAATT TTATGGTGCA CTCTCAGTAC  
 AATCTGCTCT GATGCCGCAT AGTTAAGCCA GCCCCGACAC CCGCCAACAC CCGCTGACCG  
 GCCCTGACGG GCTTGCTGCTG TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG  
 GAGCTGCATG TGTCAGAGGT TTTACCCGTC ATCACCAGAA CGCGCGAGAC GAAAGGGCCT  
 CGTGATACCG CTATTTTTAT AGGTAAATGT CATGATAATA ATGGTTTCTT AGACGTCAGG  
 TGGCACTTTT CGGGGAAATG TGCGCGGAAC CCCTATTTGT TTATTTTTCT AAATACATTC  
 AAATATGTAT CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT ATTGAAAAAG  
 GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT CCCTTTTTTG CCGCATTTTG

MI98A002509

Figura 6b

CCTTCCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG AAGATCAGTT  
GGGTGCACGA GTGGGTATACA TCGAACTGGA TCTCAACAGC GGTAAGATCC TTGAGAGTTT  
TCGCCCCGAA GAACGTTTTT CAATGATGAG CACTTTTAAA GTTCTGCTAT GTGGCGCGGT  
ATTATCCCGT ATTGACGCCG GGCAAGAGCA ACTCGGTGCG CGCATACACT ATTCTCAGAA  
TGACTTGCTT GAGTACTCAC CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG  
AGAATTATGC AGTGCTGCCA TAACCATGAG TGATAACACT GCGGCCAACT TACTTCTGAC  
AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGAC AACATGGGGG ATCATGTAAC  
TCGCCTTGAT CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG AGCGTGACAC  
CACGATGCCT GTAGCAATGG CAACAACGTT GCGCAAAC TAACCTGGCG AACTACTTAC  
TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG GATAAAGTTG CAGGACCACT  
TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG CCGGTGAGCG  
TGGGTCTCGC GGATATCATT CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT  
TATCTACACG ACGGGGAGTC AGGCAACTAT GGATGAACGA AATAGACAGA TCGCTGAGAT  
AGGTGCCTCA CTGATTAAGC ATTGGTAAC GTGAGACCA GTTTACTCAT ATATACTTTA  
GATTGATTTA AAACCTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC TTTTGTATA  
TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCGTAGA  
AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC  
AAAAAAACCA CCGCTACCAG CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT  
TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT ACTGTCCTTC TAGTGTAGCC  
GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG CTCTGCTAAT  
CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT TGGACTCAAG  
ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTTCGT GCACACAGCC  
CAGCTTGAG CGAACGACCT ACACCGAAT GAGATACCTA CAGCGTGAGC ATTGAGAAAG  
CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC  
AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG  
GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC TCGTCAGGGG GCGCGAGCCT  
ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCCCT GCCTTTTGCT GGCCTTTTGC  
TCACATGTTT TTTCTGCGT TATCCCCTGA TTCTGTGGAT AACCGTATTA CCGCCTTTGA  
GTGAGCTGAT ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG TGAGCGAGGA  
AGCGGAAGAG CGCCCAATAC GCAAACCGCC TCTCCCCGCG CGTTGGCCGA TTCATTAATG  
CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGGCAGT GAGCGCAACG CAATTAATGT  
GAGTTAGCTC ACTCATTAGG CACCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATGTT  
GTGTGGAATT GTGAGCGGAT AACAATTTCA CACAGGAAAC AGCTATGACC ATGATTACGC

C

M198A002509

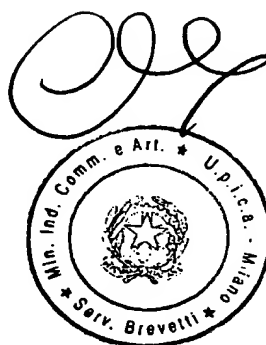
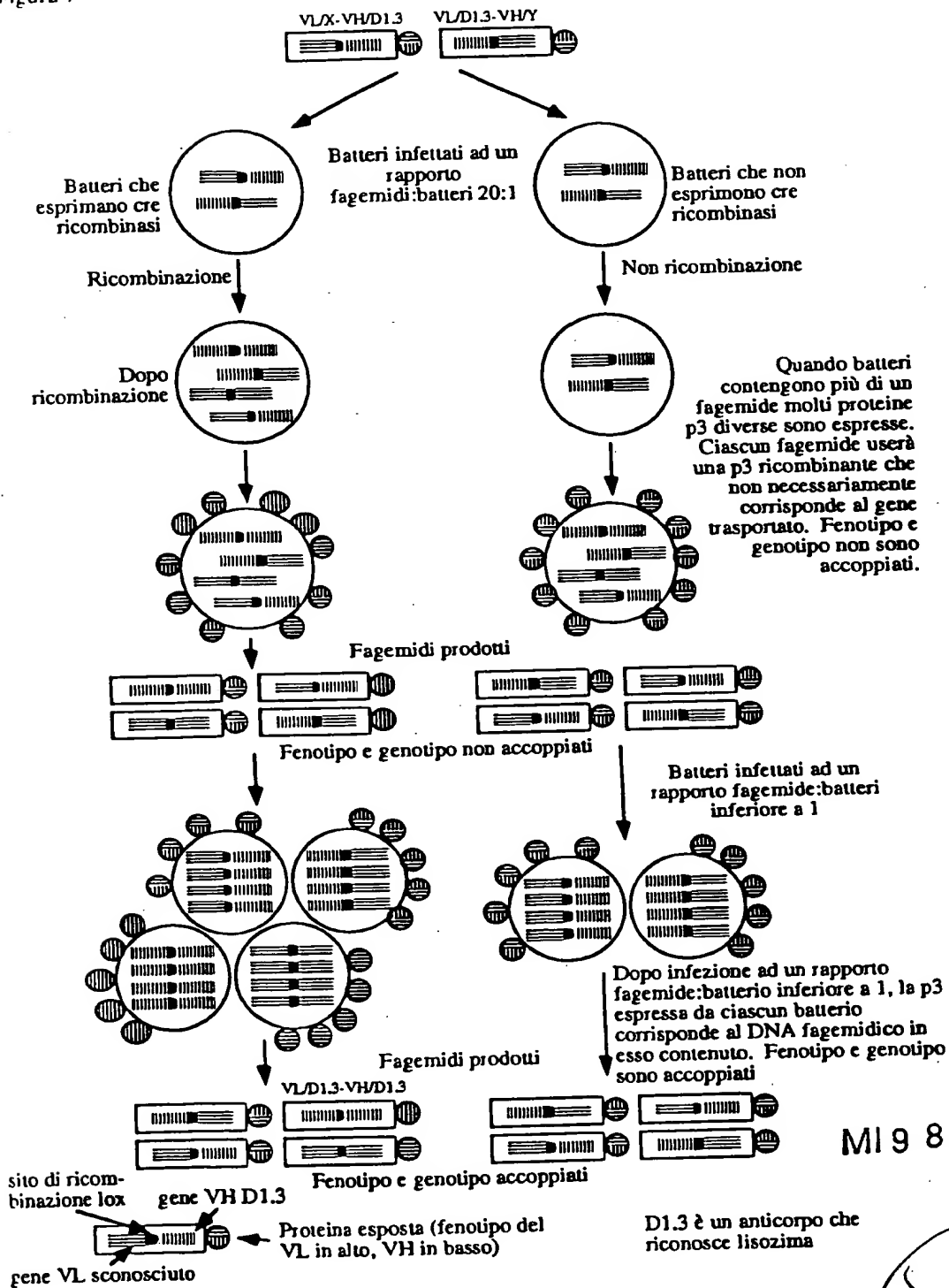
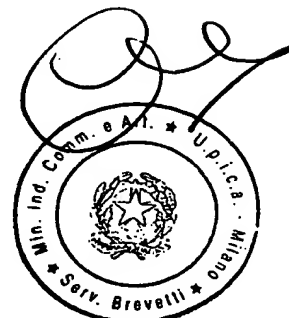


Figura 7



MI98A002509



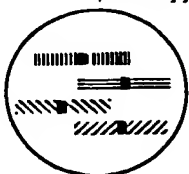
*[Handwritten signature]*

Figura 8

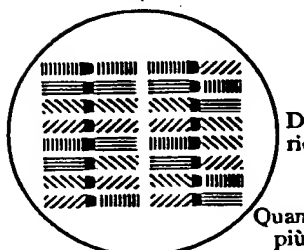
Libreria primaria ( $7 \times 10^7$ ) rappresentata da 4 fagemidi diversi



Batteri che esprimono cre ricombinasi infettati con fagemidi ad un rapporto fagemide:batterio 200:1

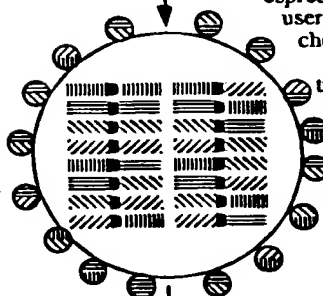


Ricombinazione

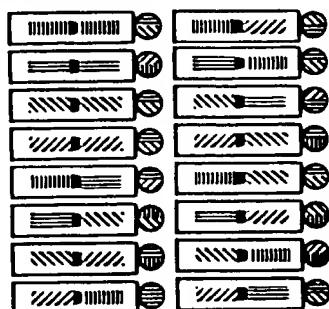


Dopo ricombinazione

Quando batteri contengono più di un fagemide molte proteine p3 diverse sono espresse. Ciascun fagemide userà una p3 ricombinante che non necessariamente corrisponde al gene trasportato. Fenotipo e genotipo non sono accoppiati.



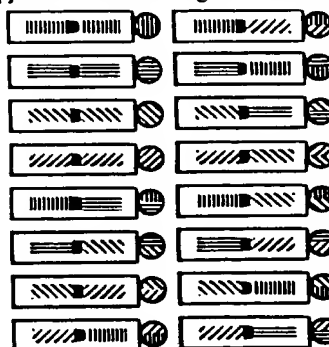
Fagemidi prodotti



Fenotipo e genotipo non accoppiati

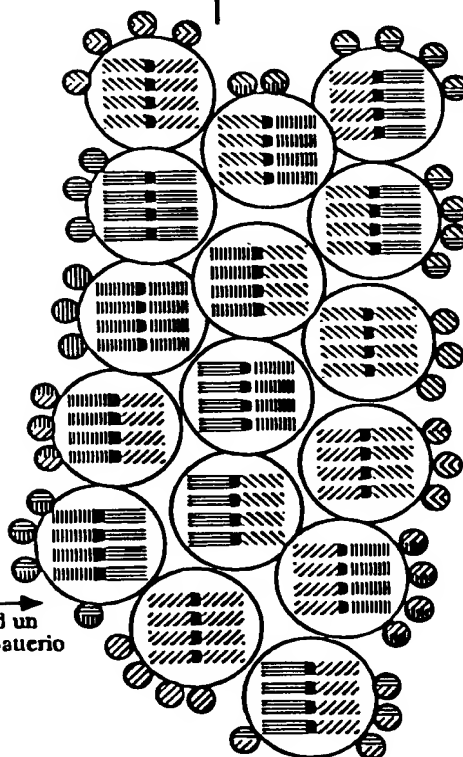
4 diversi geni  
VH/VL sconosciuti

Libreria ricombinata secondaria ( $\sim 10^{12}$ ) rappresentata da 16 fagemidi diversi



Fenotipo e genotipo accoppiati

Fagemidi prodotti



Batteri infettati ad un rapporto fagemide:batterio inferiore a 1

Dopo infezione ad un rapporto fagemide:batterio inferiore a 1, la p3 espressa da ciascun batterio corrisponde al DNA fagemidico in esso contenuto. Fenotipo e genotipo sono accoppiati. Queste colonie sono usate per stimare la diversità presente in un batterio dopo la ricombinazione.

MI98A002509

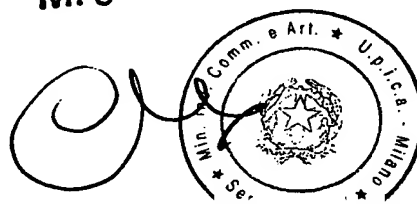




Figura 9

cellula 1

		geni VH																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
geni VL	1	1							1									
	2								1									
	3			1														
	4				1						1							
	5					1												
	6											1						
	7												1					
	8													1				
	9	2		2														
	10	3			1						1							
	11	2	1			1				1								
	12	2		1		1	1	1					2		1		1	3

17 diverse colonie VH su 37 colonie totali  
 12 diverse colonie VL su 37 colonie totali  
 28 diverse combinazioni VL-VH  
 2 scFv non hanno partecipato alla ricombinazione

cellula 2

		geni VH																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
geni VL	1	1																	
	2		1							1									
	3			1															1
	4			5	1	1	5					1			1				
	5	1					1	1	1		1							1	
	6	2																	
	7							1											
	8			3				1								1			
	9																1		
	10								1										
	11	1																	
	12												1						
	13							1											
	14			1										1					

18 diverse colonie VH su 41 colonie totali  
 14 diverse colonie VL su 41 colonie totali  
 29 diverse combinazioni VL-VH  
 2 scFv non hanno partecipato alla ricombinazione

MI98A002509

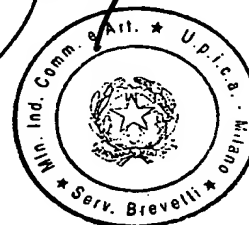
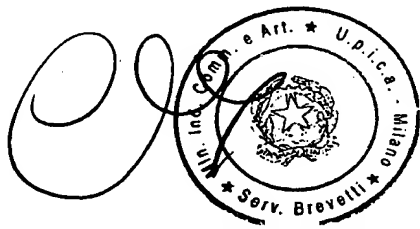
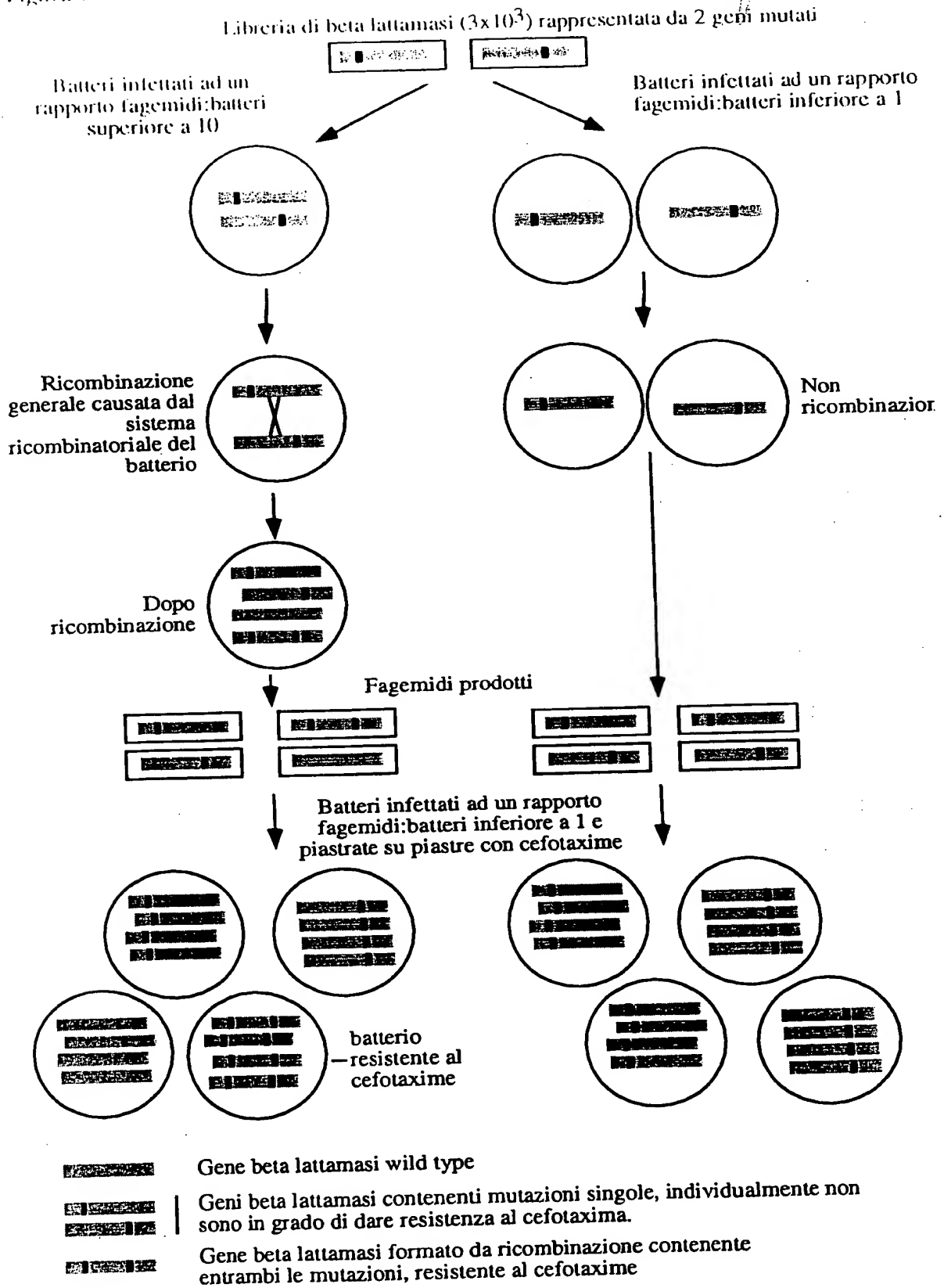


Figura 10



MI98A002509



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☐ **FADED TEXT OR DRAWING**

☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**